



Diplôme d'Etudes Supérieures Spécialisé
de Biotechnologie Végétale

Université Paul Sabatier. Toulouse III

**Cartographie génétique de gènes candidats
pour la qualité du bois
chez l'*Eucalyptus***



Audrey CARROUCHÉ

Responsable de stage :

Jean-Marc GION

Laboratoire de Génétique
Programme Arbres et Plantations
CIRAD Forêt



Remerciements

Tout d'abord, je tiens à remercier M. Mallet pour m'avoir accueillie au sein du programme Arbres et Plantations du CIRAD-forêt.

De sincères remerciements vont à également à M. Philippe Vigneron, responsable du thème Amélioration génétique, pour ses conseils et ses nombreuses explications ainsi que pour sa gentillesse.

Des remerciements tout particuliers vont à M. Jean-Marc Gion, mon maître de stage, pour m'avoir proposé ce stage, pour sa disponibilité à toute épreuve, pour la confiance et la liberté d'action accordées, pour son investissement en tant que maître de stage. Jean-Marc, 1000 fois merci.

Je n'oublierai pas non plus M. Frédéric Mortier : merci de t'être intéressé, merci d'avoir partagé un peu de cette science obscure pour moi que sont les statistiques ! Je souhaite remercier également LaTeX pour l'aide précieuse lors de la mise en forme de ce document.

Je ne sais comment remercier toutes les personnes qui ont fait vivre (et c'est un bien faible mot!!!) le laboratoire durant ces 8 mois de stage. Il est évident que ce laboratoire ne serait pas ce qu'il est sans Mireille et Alexandre, une équipe aussi différente que complémentaire. Mireille, merci d'avoir toujours été présente quand j'en avais besoin ; Alexandre, merci pour ta patience lors de mes manips SSCP, et merci pour la musique ! Et puis bien sur, mes collègues stagiaires ou thésard : Nicolas (espèce de Krotto, tu n'as pas honte de nous avoir abandonné pour aller en Equateur ? et tu trouves que c'est une bonne raison ? !), Michaël (cher KaelKael, et oui, il faut bien avouer que tu nous as manqué !! et le bruit des balles de ping-pong également !), Elodie (je suis sûre que tu vas t'éclater avec tes chèvres en Ardèche ! n'hésites pas à prendre des photos, nous en ferons bon usage !), Pierre (à quand un parpaing' sur Toulouse ? !), Lorraine (profite bien du voyage pour ceux qui restent ici !) et puis aussi les miss passées trop rapidement : Annabelle et Aurore. J'ai également une pensée particulière pour Roselyne merci d'avoir toujours été là.

Enfin, un grand merci aux Bissyen(ne)s pour avoir fait de ce séjour à Montpellier un moment agréable, merci à Carine, Jujue, Céline, Roro, Greg, et bien sur Virginie.

Sommaire

1	INTRODUCTION	1
2	MATERIEL et METHODES	7
2.1	Matériel végétal	7
2.2	Extraction d'ADN génomique	7
2.3	Caractéristiques des EST utilisées	8
2.4	Définition d'amorces spécifiques et amplification des EST	8
2.5	Révélation du polymorphisme par SSCP	9
2.5.1	Principe de la technique	9
2.5.2	Protocole utilisé	9
2.6	Cartographie Génétique	10
2.6.1	Stratégie de cartographie	10
2.6.2	Cartographie des EST	10
2.7	Caractères de qualité du bois étudiés	11
2.7.1	Description des caractères	11
2.7.2	Statistiques descriptives des variables	13
2.8	Détection de QTL	13
2.8.1	Détection marqueur par marqueur (ANOVA)	13
2.8.2	Détection par Interval Mapping (IM)	14
2.8.3	Tests de permutations	15
3	RESULTATS	16
3.1	Cartographie génétique des EST	16
3.1.1	Amplification des EST	16
3.1.2	Détection du polymorphisme par SSCP	16
3.1.3	Localisation des EST sur les cartes génétiques	18
3.2	Analyse des colocalisations EST/QTL	19
3.2.1	Variabilité des caractères quantitatifs étudiés	19
3.2.2	Analyse QTL marqueur par marqueur	20
3.2.3	Analyse QTL par Interval Mapping (IM)	21
3.2.4	Tests de permutations	22
4	DISCUSSION	23
4.1	Cartographie génétique des EST	23
4.1.1	Détection du polymorphisme	23

4.1.2	Conservation des séquences entre espèces	23
4.1.3	Informations apportées par la cartographie d'EST	24
4.2	Détection de QTL	25
4.2.1	Détection de QTL marqueur par marqueur <i>vs</i> détection par IM	25
4.2.2	Absence de colocalisation	25
4.2.3	Colocalisation des EST avec des QTL de qualité du bois . . .	26
5	CONCLUSION et PERSPECTIVES	27
	Bibliographie	28

Abréviations

- ADN Acide désoxyribonucléique
- AFLP Amplified Fragment Length Polymorphism
- ANOVA Analyse of variance
- BET Bromure d'éthidium
- BSA Bovine serum albumine
- CAPS Cleaved Amplified Polymorphic Sequence
- CCR Cinnamoyl CoA reductase
- CIRAD Centre de Coopération Internationale en Recherche Agronomique pour le Développement
- cM Centimorgan
- dNTP Désoxynucléotide triphosphate
- EDTA Acide Ethylène Diamine Tétracétique
- EST Expressed Sequence Tag
- GL Groupe de liaison
- HCl Acide chloridrique
- HSS Hybridation Soustractive Suppressive
- IM Interval Mapping
- LOD Log10 of the odds ratio
- M Molaire
- MgCl₂ Dichlorure de magnésium min Minute
- ml Millilitre
- mM Millimolaire
- NaOH Hydroxyde de sodium
- ng Nanogramme
- pb Paire de base
- PCR Polymerase Chain Reaction
- QTL Quantitative Trait Loci
- RAPD Random Amplified Polymorphic DNA
- RFLP Restriction Fragment Length Polymorphism
- SAM Sélection assistée par marqueurs
- SCE Somme des carrés des écarts
- SSCP Single Strand Conformation Polymorphism
- TBE Tris borate EDTA
- UV Ultraviolet
- V Volt
- μ g micro gramme
- μ l micro litre
- μ mol micro litre

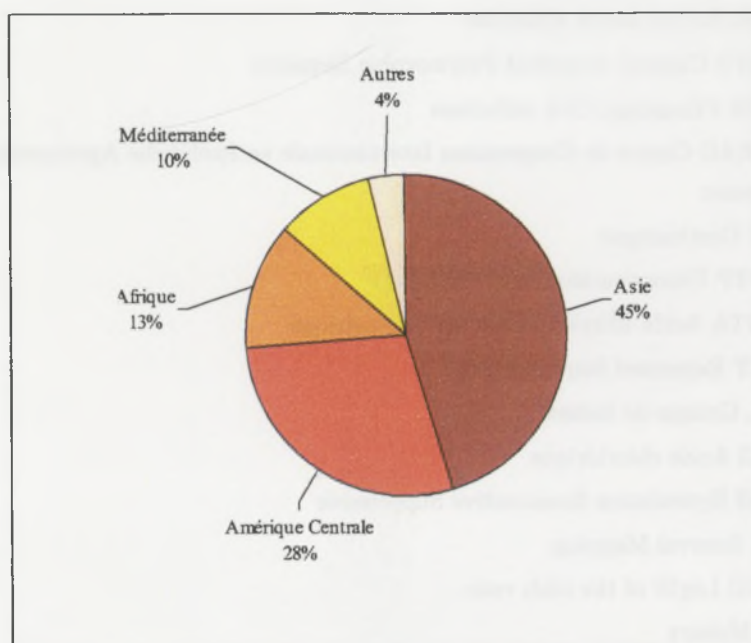


Figure 1 : Répartition géographique des 13 millions d'hectares de plantations d'eucalyptus dans le monde.

1 INTRODUCTION

L'eucalyptus est un des arbres forestiers les plus utilisés en reboisement. En 2001, 13,7 millions d'hectares de plantations ont été recensés à travers le monde (FAO, 2001 [16]), dont près de la moitié en Asie (Figure 1). Sur les 600 espèces que compte le genre *Eucalyptus*, seules une dizaine sont exploitées dans les régions d'introduction dont *E. urophylla*, *E. grandis* et *E. globulus*. Les raisons de cette utilisation sont d'abord liées aux propriétés sylvicoles intéressantes en plantation : une adaptation à de nombreux milieux, une forte croissance au jeune âge qui permet une exploitation rapide du bois (âge de rotation de 7 ans, contre 40 ans chez les pins et 20 ans chez les peupliers) et enfin une bonne aptitude au bouturage qui permet une propagation clonale des meilleurs génotypes. De plus, le bois d'eucalyptus est sujet à de nombreuses utilisations. D'après Vigneron et Bouvet (1997 [60]), 80% de la production de bois d'eucalyptus sert comme bois d'énergie (bois de chauffe et charbon de bois). Le bois de trituration (panneaux de particules et pâte à papier) représente aussi une part importante de l'utilisation des eucalyptus notamment celui issu des plantations industrielles. En Australie, aire d'origine du genre, il est utilisé comme bois d'oeuvre, mais cette utilisation reste rare dans les pays d'introduction. En effet, le bois issu de plantations peut présenter à la coupe et au séchage de nombreux défauts : retrait, fentes en bout... D'autres valorisations non ligneuses peuvent également être citées comme l'extraction d'huiles essentielles (Moran et Bell, 1983 [37]) pour la pharmacopée, la production de miel...

L'accroissement de la productivité des plantations forestières s'est faite progressivement par l'amélioration des techniques sylvicoles mais aussi par le développement de programmes d'amélioration génétique. L'amélioration génétique des arbres forestiers est encore très récente et ne concerne qu'un nombre restreint d'espèces de pins, de peupliers, de sapins, d'eucalyptus (Plomion, 2000 [43]). Même si la qualité du bois est un caractère économique majeur, ce caractère complexe reste encore peu pris en compte dans ces programmes d'amélioration par rapport aux caractères de croissance plus faciles et moins coûteux à évaluer.

Pour l'eucalyptus, un programme d'amélioration génétique a vu le jour au Congo dans les années 50 avec l'introduction de plusieurs espèces. Suite à ces introductions, des hybrides naturels sont apparus, caractérisés par une croissance et une adaptation supérieures aux espèces introduites. La maîtrise du bouturage herbacé (Martin et Quillet, 1974 [36]) a permis la multiplication à grande échelle de ces hybrides.

Ensuite, le développement de la pollinisation contrôlée (Maillard, 1978 [34]) a permis de tester différentes combinaisons hybrides. Plusieurs de ces hybrides artificiels se sont révélés plus intéressants que les hybrides naturels, notamment les croisements interspécifiques *E. urophylla* x *E. pellita* et *E. urophylla* x *E. grandis*. Depuis 1989, un schéma de sélection récurrente réciproque (SRR) a été développé pour le croisement *E. urophylla* x *E. grandis* (Vigneron, 1989 [59]). Le principe de la SRR est d'améliorer conjointement deux groupes d'individus l'un par rapport à l'autre, dans ce cas deux espèces, de manière à obtenir des hybrides inter-groupes (interspécifiques) regroupant certains caractères spécifiques à chacun des deux groupes. En effet *E. urophylla* présente une bonne adaptation aux conditions congolaises alors qu'*E. grandis* est mal adapté mais possède une capacité de croissance importante.

Les programmes d'amélioration des arbres forestiers sont coûteux en terme d'espace et de temps. De plus, chez ces espèces, il existe souvent une mauvaise corrélation juvénile-adulte pour beaucoup de caractères d'intérêts économiques (Bouvet, 1995 [4]). L'évaluation de ces caractères ne peut donc se faire que tardivement (3 ans au plus tôt chez l'eucalyptus au Congo). Ensuite, la maturation sexuelle tardive (3 à 5 ans pour l'eucalyptus) limite les possibilités de croisements. Enfin, la plupart des caractères économiques importants (comme la qualité du bois et la croissance) sont des caractères complexes à héritabilité faible à moyenne, variant de 0,1 à 0,4 (Kremer, 1994 [30]). En effet, ces caractères ont souvent un déterminisme polygénique et sont fortement soumis aux effets du milieu. L'efficacité de la sélection n'est donc pas optimale.

Dans le but de raccourcir le temps d'obtention de nouvelles variétés et d'augmenter l'efficacité de la sélection, deux stratégies sont envisageables (Plomion, 2000 [43]). Tout d'abord, la création de géotypes intéressants par transformation génétique, multipliés par embryogenèse somatique (Harvengt *et al.*, 2000 [27]). Mais chez l'eucalyptus, l'embryogenèse somatique n'est pas encore utilisable du fait de difficultés de régénération. De plus, cette proposition pose le problème de l'opinion publique face aux OGM. La deuxième solution passe par l'utilisation de marqueurs moléculaires dans les programmes de sélection (sélection assistée par marqueurs ou SAM). Elle s'appuie sur l'exploration et la mise à profit de la variabilité génétique naturelle. Plusieurs travaux sont menés afin d'utiliser l'information moléculaire pour faciliter la prise en compte de la qualité du bois dans les programmes d'amélioration classique.

La qualité du bois est un caractère complexe et hétérogène, variant entre génotypes mais également au sein même de l'arbre. Classiquement, on distingue 6 types de bois aux propriétés chimiques, anatomiques, physiques et mécaniques différentes : bois juvénile, bois mature, bois de réaction, bois normal, bois initial et bois final. Chez les eucalyptus du Congo, seuls les 4 premiers types de bois sont rencontrés, en effet, du fait de la croissance continue des arbres, la distinction bois initial bois final n'est pas observée. Dans le cadre des angiospermes, le bois de tension (bois de réaction) est un bois qui peut être induit par une déstabilisation des conditions de croissance (vent. . .) ou par une réorientation traumatique (mort d'un apex)(Baillères, 1994 [2]). Selon l'utilisation du bois, les propriétés recherchées sont différentes. Ainsi, lorsque le bois est utilisé en trituration, les paramètres importants sont la qualité des fibres de cellulose et la composition chimique du bois (la teneur en lignines doit être diminuée tandis que celle de cellulose doit être augmentée). Par contre, pour le bois d'énergie, une teneur plus élevée en lignine sera souhaitée car son pouvoir calorifique est plus élevé que celui de la cellulose. Enfin, pour le bois d'oeuvre, ce sont les propriétés mécaniques et physiques qui sont importantes, notamment la faible nervosité du bois. En effet, un bois nerveux libère lors de la coupe les contraintes de croissance qui vont se traduire par des déformations des pièces de bois et principalement par l'apparition de fentes en bout et de retraits au séchage.

Ces différentes caractéristiques macroscopiques sont dues aux propriétés intrinsèques du bois, c'est à dire à la structure et la composition des parois secondaires des cellules du xylème mature (bois). Le xylème est le tissu de soutien mécanique servant à la conduction de sève brute des racines aux feuilles. Il est composé de fibres, de cellules parenchymateuses et de trachéides. La maturation des cellules du xylème se fait en 4 étapes : expansion cellulaire, mise en place de la paroi secondaire, lignification et enfin mort cellulaire (Plomion *et al.*, 2001 [48]). Durant la phase d'expansion cellulaire, les cellules s'agrandissent longitudinalement et radialement. Lorsque la paroi primaire est formée, ces cellules atteignent leur taille finale. Puis la paroi secondaire se met en place par l'assemblage de ses 4 composés majeurs : les polysaccharides (celluloses et hémicelluloses), la lignine, les protéines de la paroi et des composés mineurs solubles et insolubles. La cellulose représente 40 à 50% du bois. Elle est constituée de chaînes de β -glucoses associés par des liaisons 1-4. De nombreuses liaisons intra et inter moléculaires permettent un assemblage rigide en micro fibrilles (Fukuda, 1996 [18]). Les hémicelluloses constituent 25% du poids sec du bois. Elles sont constituées d'un axe osidique auquel sont rattachées des chaînes de résidus et

sont fortement liées aux micro fibrilles de cellulose. La lignine est le troisième composant majeur du bois et en représente 25 à 35%. Les lignines sont des polyphénols qui sont obtenus par polymérisation de 3 monolignols : le para-hydroxyphényl (unité H), le guaiacyl (unité G) et le syringyl (unité S). C'est une des voies de biosynthèse les mieux connues et les plus étudiées (Grima-Pettenati et Goffner, 1999 [26]). Une fois la lignification terminée, un programme de mort cellulaire se met en place. Des hydrolases dégradent alors tout le contenu de la cellule, tout en maintenant la paroi secondaire lignifiée (Fukuda, 1996 [18]).

La xylogénèse est un processus complexe faisant intervenir un nombre important de voies métaboliques. Dans un but d'amélioration génétique pour la qualité du bois, il apparaît nécessaire de mieux connaître les mécanismes moléculaires qui gouvernent ce processus. Aujourd'hui, de nombreux travaux vont dans ce sens, notamment grâce aux avancées technologiques en biologie moléculaire.

En effet, des programmes de séquençage total de génomes d'arbres forestiers sont en cours (peuplier, eucalyptus) ainsi que plusieurs autres pour le génome exprimé (banques d'EST développées chez de nombreuses espèces : pins, peupliers, eucalyptus, bouleau, chêne...). Un grand nombre de ces banques d'EST est développé pour l'étude de l'expression des gènes impliqués dans la xylogénèse. Chez le pin, Allona *et al.* (1998 [1]) ont montré que 55% des EST obtenues à partir de xylème immature présentaient des similarités avec des gènes de fonction connue ou avec des séquences disponibles dans les banques de données. Parmi ces EST, 10% codent des protéines impliquées dans la formation des parois (enzymes de la lignification, protéines pariétales comme les extensines...). Des gènes impliqués dans le métabolisme des acides aminés et la régulation du métabolisme cellulaire ont également été mis en évidence. Chez le peuplier, des résultats similaires ont été obtenus pour des banques d'EST réalisées à partir de la région cambiale et de xylème (Sterky *et al.*, 1998 [53]). Dans cette étude, un autre transcrit abondant a été mis en évidence : la S-Adenosyl-L-méthionine qui joue un rôle dans la méthylation des précurseurs des monolignols. Ces études d'expression de gènes sont souvent réalisées grâce aux transcrits pour lesquels différentes techniques sont disponibles. Certaines d'entre elles permettent d'obtenir des banques d'EST enrichies pour une condition donnée. Une première technique dite d'AFLP-ADNc a été utilisée récemment chez le pin maritime pour le développement d'une banque d'EST à partir de xylème en différenciation (Leprovost, 2003 [33]).

Diatchenko *et al.* (1996 [15]) ont développé la technique dite d'Hybridation Sous-tractive Suppressive (HSS). Cette technique a été utilisée chez *E. gunnii* pour obtenir une banque EST enrichie en gènes surexprimés dans le xylème par rapport à la feuille (Paux, Thèse en cours). La HSS est de plus en plus utilisée pour obtenir des banques enrichies, à titre d'exemple, pour l'étude des gènes exprimés lors de stress chez *Pisum sativum* (Sävenstrand *et al.*, 2000 [54]). Ces banques d'EST permettent d'étudier l'expression d'un grand nombre de gènes dans différents organes, différents tissus ou encore pour plusieurs conditions environnementales. Chez le peuplier, une étude d'expression des gènes a été réalisée pour différents stades de la xylogénèse à partir de 3000 EST (Hertzberg *et al.*, 2001 [29]). Une analyse micro array a permis de montrer que la régulation de la transcription pour différents gènes, notamment ceux codant les enzymes de biosynthèse des lignines et de la cellulose, dépendait du stade de développement du xylème.

Cependant, pour que ces gènes puissent être utilisés en sélection, il est nécessaire que leur variabilité moléculaire en population soit reliée à la variation du caractère d'intérêt ciblé. Dans cette optique, plusieurs travaux de cartographie génétique et de détection de QTL ont été développés chez les arbres forestiers. Les premières espèces forestières pour lesquelles ces types de travaux ont été réalisés sont les pins (Plomion *et al.*, 1995 [44]), le peuplier (Bradshaw *et al.*, 1995 [6]; Cervera *et al.*, 1999 [11]) et différentes espèces d'*Eucalyptus* (Grattapaglia et Sederoff, 1994 [25]; Byrne *et al.*, 1995 [9]; Verhaegen et Plomion, 1996 [57]; Brondani *et al.*, 1998 [8]). Ces cartes génétiques ont été obtenues principalement grâce à des marqueurs anonymes dominants (de type RAPD et AFLP) et/ou codominants (de type RFLP et SSR). Grâce à ses cartes génétiques, des QTL (Quantitative Trait Loci) ont pu être mis en évidence pour différents caractères : la croissance chez l'eucalyptus (Grattapaglia *et al.*, 1996 [23]; Verhaegen *et al.*, 1997 [58]; Gion, 2001 [21]), le pin (Plomion *et al.*, 1996 [46, 45]), le peuplier (Bradshaw et Stettler, 1995 [5]), la réponse au stress environnemental et la propagation végétative chez l'eucalyptus (Byrne *et al.*, 1997 [10]; Grattapaglia *et al.*, 1995 [24]). Enfin, des QTL ont aussi été détectés pour la qualité du bois chez le pin (Neale *et al.*, 1997 [39]) et l'eucalyptus (Verhaegen *et al.*, 1997 [58]). Cependant, l'utilisation de marqueurs anonymes limite considérablement ce type d'approche dans l'optique d'une utilisation en SAM.

Depuis peu, des stratégies plus ambitieuses ont été développées dites approches gène candidat. L'idée est d'utiliser des gènes de fonction connue qui pourraient expliquer *a priori* la variation du caractère étudié. Jusqu'à aujourd'hui, peu de travaux

rapportent des colocalisations gènes/QTL chez les arbres forestiers. Pour des gènes de la voie de biosynthèse des lignines, positionnés sur les cartes génétiques d'*E. urophylla* et d'*E. grandis* (Gion *et al.*, 2000 [22]), des colocalisations ont pu être observées entre des gènes et des QTL de teneur en lignines (Gion, 2001 [21]). Des résultats similaires ont été observés chez le peuplier. Frewen *et al.* (2000 [17]) ont mis en évidence des colocalisations entre des gènes impliqués dans la perception de la photopériode et dans la transduction des signaux de réponse à l'acide abscissique et des QTL affectant la levée de dormance des bourgeons.

Pour révéler du polymorphisme au sein d'un gène candidat, plusieurs techniques peuvent être employées. Ainsi, des cartes génétiques ont été construites à partir d'EST et de gènes connus grâce à des marqueurs RFLP (polymorphisme de longueur de fragments de restriction) chez *E. globulus* (Thamarus *et al.*, 2002 [55]). D'autres techniques peuvent également être utilisées : DGGE (électrophorèse de l'ADN en gradient dénaturant ; Myers *et al.*, 1985 [38]) et TGGE (électrophorèse de l'ADN en gradient de température ; Riesner *et al.*, 1992 [50]). Ces deux techniques sont basées sur des différences de stabilité de l'ADN. Les CAPS (Cleaved Amplified Polymorphic Sequence ; Tragoonrun *et al.*, 1992 [56]) ont également été utilisés pour cartographier des gènes candidats pour la résistance au virus de la mosaïque de la canne à sucre chez le maïs (Quint *et al.*, 2002 [49]). Une dernière technique, peu coûteuse, est la SSCP (single-strand conformation polymorphism ou polymorphisme de conformation du simple brin). Cette technique a été développée par Orita *et al.* (1989 [41]) pour la détection de mutations de l'ADN génomique. Elle a ensuite été utilisée pour faire de la cartographie génétique de gènes de fonction connue chez le pin (Plomion *et al.*, 1999 [47]) et l'eucalyptus (Gion *et al.*, 2000 [22]).

Mon stage s'inscrit dans cette approche gène candidat pour la qualité du bois, menée par le Cirad-Forêt. L'objectif est de cartographier des EST, obtenues par HSS chez *E. gunnii*, qui sont surexprimées dans le xylème par rapport à la feuille. Pour cela, il a fallu définir des amorces spécifiques pour les différentes EST étudiées puis révéler du polymorphisme au sein des fragments amplifiés par SSCP. Ensuite, les EST polymorphes ont été cartographiées sur les cartes génétiques déjà existantes afin d'étudier leur co-localisation avec des QTL de qualité du bois.

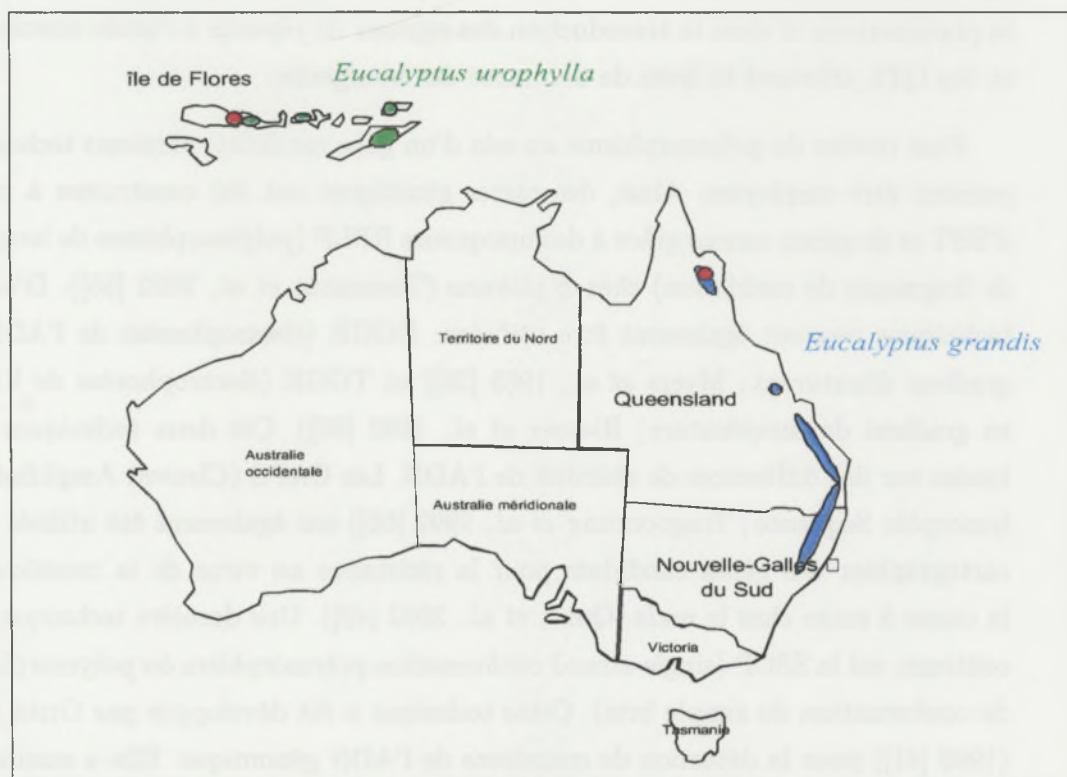


Figure 2 : Aire naturelle des deux espèces d'*Eucalyptus* utilisées. La localisation des provenances des deux géniteurs est indiquée en rouge.

2 MATERIEL et METHODES

2.1 Matériel végétal

La descendance de cartographie utilisée correspond à un croisement interspécifique entre *E. urophylla* et *E. grandis*. Le géniteur *E. urophylla* (individu 14-144), utilisé comme parent femelle, provient du Mont Lewotobi de l'Île de Florès tandis que le géniteur *E. grandis* (individu 9-21) provient d'Atherton en Australie. Les aires d'origine de ces 2 espèces sont représentées sur la figure 2, ainsi que les provenances des 2 géniteurs. Cette descendance élite, plantée au Congo en 1992, totalise 201 individus pleins-frères et correspond au croisement de référence pour les différents travaux de cartographie et de détection de QTL réalisés jusqu'à maintenant (Verhaegen et Plomion, 1996 [57]; Verhaegen *et al.*, 1997 [58]; Gion *et al.*, 2000 [22]). Pour cette étude, un échantillon de 92 individus a été utilisé sur les 200 disponibles.

2.2 Extraction d'ADN génomique

L'extraction d'ADN génomique a été réalisée selon un protocole modifié de Saghai-Maroo *et al.* (1984 [52]) à partir de jeunes feuilles séchées, prélevées sur des arbres adultes. Pour chaque individu, 100mg de feuilles séchées sont broyées en présence d'azote liquide jusqu'à obtention d'une fine poudre. Cette poudre est incubée 1h à 65°C en présence de 1,3ml de tampon d'extraction stérile (Tris HCl 0,1M pH8, EDTA 0,02M, ATMA 20g/L, NaCl 1,4M), de 0,1% de mercaptoéthanol et de 1% de polyvinylpyrrolidone. Ensuite, 400µl de dichlorométhane sont rajoutés et les extraits sont mis à agitation douce pendant 10min puis centrifugés 10min à 6000T/min. Le surnageant est prélevé et cette étape est renouvelée 2 fois. L'ADN génomique contenu dans le surnageant est précipité à froid par l'ajout de 500µl d'isopropanol glacé et de 7µl de NaCl 5M. Les extraits sont placés 1h à -20°C puis centrifugés 10min à 6000T/min. Les culots sont ensuite lavés par 500µl d'éthanol à 75°, séchés sous vide à 40°C et remis en suspension dans 200 à 500µl d'H₂O stérile, suivant la taille du culot.

L'ADN est quantifié avec un fluorimètre (Fluoroscanner). Pour chaque extrait, 2µl d'ADN génomique sont ajoutés à 200µl de mélange réactionnel [TNE 1X (SigmaTM), HOESCHST (bis benzimide) à 0,1µg/ml]. Une gamme étalon de 0 à 300ng/µl est réalisée à partir d'ADN de thymus de veau. Les échantillons sont ensuite dilués avec de l'eau stérile à 3ng/µl. Ces solutions de travail sont conservées à -20°C.

Tableau 1 : Description des EST utilisées pour l'approche gène candidat fonctionnel. La fonction probable des EST a été obtenue grâce au logiciel BLASTX contre la banque NR. X/F désigne le niveau d'expression de l'EST dans le xylème par rapport à la feuille chez *E. gunnii*.

EST Id.	X/F	Taille en pb	FONCTION	SCORE	E-value
06A01	1,8	523	putative protein [Arabidopsis thaliana]	38,9	0,041
06A05	4,5	411	reverse transcriptase [Cicer arietinum]	36,6	0,073
06A07	8,1	493	dehydrin [Phaseolus vulgaris]	47,8	7E-05
06A08	7,5	507	17,7 kDa heat shock protein [Helianthus annuus]	110	1E-23
06A11	1,9	540	xyloglucan endo-1,4-beta-D-glucanase [Arabidopsis thaliana]	247	7E-65
06A12	1,3	248	glycosyl hydrolase family 10 [Arabidopsis thaliana]	92,4	1E-18
06B01	2,1	602	Homeobox protein SBH1 [Glycine max]	328	2E-89
06B07	3,8	325	putative protein [Arabidopsis thaliana]	62,0	2E-09
06B08	4,9	311	No Hits		
06B10	6,1	517	No Hits		
06C03	1,6	266	limonoid UDP-glucosyltransferase [Citrus unshiu]	59,3	1E-08
06C08	2,1	168	No Hits		
06D01	4,2	491	OSJNBa0072F16,18 [Oryza sativa]	45,4	3E-07
06D02	4,9	264	apoptosis inhibitor hiap-2 - human	32,0	2,2
06D03	4,4	433	ribosomal protein [Arabidopsis thaliana]	134	4E-31
06D07	2,9	283	unknown [Arabidopsis thaliana]	154	4E-37
06D12	3,0	246	protein serine/threonine kinase [Arabidopsis thaliana]	40,0	6E-08
06F03	1,4	446	RAC1 [Lotus japonicus]	240	3E-63
06F06	1,9	272	putative zinc finger protein [Arabidopsis thaliana]	157	5E-38
06G03	1,7	361	dehydrin - kidney bean	31,2	3,3
06G04	1,9	263	pectate lyase [Zinnia elegans]	88,2	3E-17
06H01	2,9	259	1, 4-alpha-glucan branching enzyme [Chlamydia trachomatis]	30,0	8,5
06H12	2,8	326	histone H2A [Euphorbia esula]	100	5E-21
07A01	3,0	536	cobalamine-independent methionine synthase [Solenostemon scutellarioides]	199	2E-50
07A06	1,3	456	GH1 protein - soybean	84,3	6E-16
07A07	1,8	663	xyloglucan endotransglycosylase XET1 [Asparagus officinalis]	221	9E-57
07A08	2,1	262	putative phytochelatin synthetase [Arabidopsis thaliana]	141	2E-33
07A09	0,8	212	pectinacetyltransferase [Arabidopsis thaliana]	57,4	5E-08
09G11	7,2	515	No Hits		
11E07	4,2	560	glycosyl hydrolase family 17 [Arabidopsis thaliana]	137	7E-32
11H04	4,2	624	dTDP-glucose 4,6-dehydratase, putative [Arabidopsis thaliana]	129	2E-29
12A02	2,1	522	putative aquaporin PIP2-2 [Vitis berlandieri x Vitis rupestris]	248	2E-65
12A06	1,4	382	xyloglucan endotransglycosylase [Arabidopsis thaliana]	182	9E-46
12A09	2,4	346	No Hits		

2.3 Caractéristiques des EST utilisées

Les EST utilisées correspondent à des séquences de gènes d'*E. gunnii* préférentiellement exprimés dans le xylème par rapport à la feuille, obtenues par HSS. Pour les 34 EST étudiées, la fonction, la taille et leur niveau d'expression dans le xylème sont précisés dans le tableau 1. Ces EST ont été développées chez *E. gunnii*, elles correspondent en majorité à des extrémités 3' de gènes et leur taille est comprise entre 168 et 663 pb. La plupart sont impliquées directement ou non dans la synthèse des parois, d'autres jouent un rôle dans le transport de l'eau, d'autres encore ont des fonctions "classiques" (exemple de l'EST 6D03 qui code une protéine ribosomale), enfin, certaines n'ont pas de fonction connue mais des niveau d'expression dans le xylème très importants.

2.4 Définition d'amorces spécifiques et amplification des EST

Pour chaque EST, une paire d'amorces spécifiques à la séquence d'*E. gunnii*, a été définie avec le logiciel OLIGO 4.0 (Rychlik *et al.*, 1990 [51]). Des paramètres thermodynamique liés à la séquence des amorces ont été pris en compte pour leur définition (différence de température d'hybridation entre les 2 amorces inférieure à 2°C, pourcentage en GC de 35% minimum et formation de dimères avec une énergie supérieure à -7kcal/mol). La taille des amorces est d'environ 20 nucléotides. La taille du fragment à amplifier a été aussi un critère important pour la définition des amorces. Afin de faciliter la mise en évidence de polymorphisme, les amorces ont été choisies pour amplifier des fragments au minimum de 100 pb.

Les réactions PCR sont réalisées sur un Mastercycler Gradient Eppendorf en micro plaques de 96 puits dans un volume final de 25µl avec 15ng d'ADN génomique, 0,2µM de chaque amorce, 0,04U de Taq DNA polymérase, 0,2mM de chaque dNTP, 0,4mg/ml de BSA, 0,05% de détergent W1, 12,5µl de tampon 10X stérile [166mM (NH₄)₂SO₄, 670mM tris HCl pH=8, 20mM MgCl₂, 0,7% mercaptoéthanol]. Les conditions d'amplification sont les suivantes : une dénaturation initiale de 4 min à 94°C, suivie de 35 cycles comprenant une dénaturation (30 sec à 94°C), une hybridation (1 min à une température spécifique pour chaque EST) et une élongation (1 min à 72°C) et enfin, une élongation finale de 5 min à 72°C. La température d'hybridation spécifique à chaque couple d'amorces a été déterminée par des tests de gradient de température allant de 48 à 68,5°C.

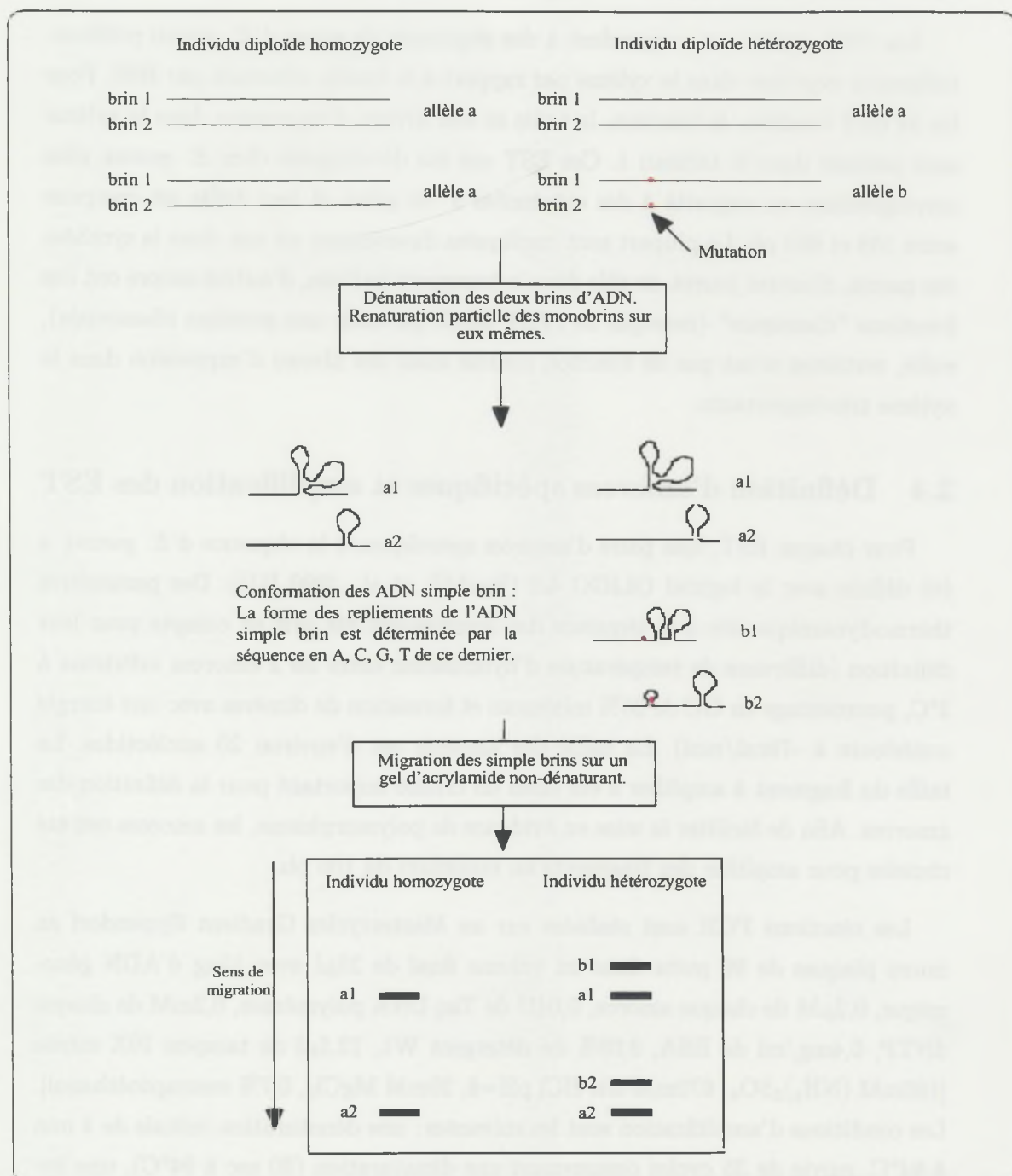


Figure 3 : Révélation du polymorphisme par la technique SSCP (Single Strand Conformation Polymorphism).

Pour chaque produit PCR, la spécificité de l'amplification est vérifiée après une électrophorèse sur gel d'agarose à 2% dans du TBE 0,5X et une coloration au bromure d'éthidium (BET).

2.5 Révélation du polymorphisme par SSCP

2.5.1 Principe de la technique

La technique SSCP (Single Strand Conformation Polymorphism), développée par Orita *et al.* (1989 [41]), est basée sur le comportement électrophorétique d'un fragment d'ADN simple brin dans un gel d'acrylamide en conditions non dénaturantes. Son principe, illustré sur la figure 3, est d'utiliser la formation de structures secondaires de simples brins d'ADN due à des appariements entre bases, pour détecter des changements de séquences au sein d'un fragment d'ADN. Ces structures secondaires dépendent des séquences nucléotidiques des simples brins et confèrent à chaque fragment une conformation particulière.

2.5.2 Protocole utilisé

Le protocole SSCP utilisé est celui décrit par Bodénès *et al.* (1996 [3]). Brièvement, 5 μ l de produits PCR sont ajoutés à 10 μ l d'une solution contenant 95% de formamide, 10mM de NaOH, 0,05% de vert de xylène cyanol et 0,05% de bleu de bromophénol. Les échantillons sont dénaturés 4 min à 94°C puis immédiatement refroidis dans la glace. Ils sont ensuite chargés sur un gel d'acrylamide (acrylamide/bisacrylamide : 37,5/1 ; 0,6X TBE ; 0,75mm x 16cm x 18cm) non dénaturant pour une migration à voltage et température constants. Les conditions de migration (pourcentage d'acrylamide du gel, température du tampon de cuve, durée de la migration, voltage) ont été optimisées pour chaque EST.

La coloration des gels au nitrate d'argent a été réalisée selon la méthode de Creste *et al.* (2001 [14]). Les gels sont traités par une succession de bains : acide acétique et éthanol, acide nitrique, nitrate d'argent, carbonate de sodium et pour finir, acide acétique. Les gels sont ensuite séchés sous vide à 80°C entre deux feuilles de papier cellophane.

2.6 Cartographie Génétique

2.6.1 Stratégie de cartographie

Afin d'intégrer les marqueurs SSCP codominants aux cartes construites avec des marqueurs dominants, ils sont codés comme des marqueurs dominants, c'est à dire que pour chaque allèle de l'EST qui ségrège dans la descendance, sa présence chez un individu est notée 1 tandis que son absence est notée 0. Lorsqu'un seul des parents est hétérozygote, la ségrégation observée doit correspondre à une ségrégation mendélienne de type 1/2 : 1/2. Dans ce cas, la cartographie de l'EST sera possible sur la carte génétique du parent hétérozygote. Lorsque les 2 parents sont hétérozygotes, la ségrégation mendélienne attendue est de type : 1/4 : 1/4 : 1/4 : 1/4. Le marqueur pourra alors être cartographié chez les 2 parents indépendamment et permettra d'identifier les groupes de liaison homologues entre les deux cartes génétiques parentales.

Pour chaque EST et chaque type de ségrégation, un test du χ^2 ($\alpha = 5\%$) a été réalisé pour vérifier que les ségrégations observées sont conformes avec une ségrégation mendélienne.

2.6.2 Cartographie des EST

On estime que 2 marqueurs sont liés génétiquement lorsqu'on observe une coségrégation entre ces deux marqueurs. Les marqueurs sont donc soumis deux à deux à un test d'indépendance de ségrégation, ou test de liaison, pour lequel l'hypothèse nulle est $H_0 : \theta = 0,5$ c'est à dire qu'il n'y a pas de liaison significative entre les marqueurs (θ étant le taux de recombinaison, compris entre 0 et 0,5). Le test de liaison est réalisé par le calcul d'un rapport de vraisemblance ou LOD score défini par :

$$LOD = \log_{10} \left[\frac{L(\theta)}{L(1/2)} \right]$$

où $L(\theta)$ est la vraisemblance d'un taux de recombinaison compris entre 0 et 0,5 et $L(1/2)$ est la vraisemblance d'une absence de liaison entre les 2 marqueurs. Pour positionner les EST sur les groupes de liaison, un LOD seuil de 4 et un θ de 0,30 ont été choisis. Ainsi, la probabilité de liaison entre marqueurs d'un même groupe est 10^4 fois supérieure à celle de l'indépendance entre marqueurs. Pour vérifier que l'ordre des marqueurs est optimum, une fonction d'ordonnement local a été utilisée. Elle réalise des permutations des triplets de marqueurs adjacents et permet d'identifier l'ordonnement le plus probable. Pour déterminer les distances génétiques entre

les différents marqueurs, la distance de Kosambi a été choisie car elle tient compte des phénomènes d'interférence, c'est à dire des contraintes mécaniques lors de la méiose qui empêchent la formation de double crossing-over entre 2 marqueurs, et ce d'autant plus que la distance entre ces 2 marqueurs est faible.

Le positionnement des EST polymorphes sur les cartes génétiques de chaque géniteur a été réalisé avec le logiciel MAPMAKER v2.0 (Lander *et al.*, 1987 [31]).

2.7 Caractères de qualité du bois étudiés

L'étude de la qualité du bois est réalisée sur différents types de caractères : chimiques, mécaniques, physiques ou encore anatomiques. Les mesures de ces caractères ont été effectuées dans le cadre des projets BIOTECH et INRA-CIRAD 2001-2003.

2.7.1 Description des caractères

- Caractères chimiques :

Deux types de caractéristiques chimiques du bois ont été considérés : les teneurs en lignines (quantitative et qualitative) et celles en flavonoïdes. La composition en lignines est estimée par deux mesures : (i) le taux de lignine, mesuré par hydrolyse en milieu acide sulfurique concentré de la cellulose et des hémicelluloses et précipitation des lignines. Le résidu insoluble obtenu, filtré, lavé et séché, correspond à la lignine de Klason, exprimée en pourcentage de résidu pariétal. (ii) la composition monomérique obtenue selon le protocole de Lapierre *et al.* (1986 [32]). La lignine est dépolymérisée par thioacydolyse et les résidus S, G, H sont analysés par chromatographie en phase gazeuse. Pour notre étude, nous avons utilisé les teneurs en S et G ainsi que S+G+H et enfin le rapport S/G. La composition en flavonoïdes a été estimée d'une part grâce au dosage des phénols réalisé par le réactif de Folin (les résultats sont exprimés en μg équivalent acide gallique par gramme de matière sèche et le caractère est noté estfol) ; d'autre part grâce au dosage des flavanols réalisé par le réactif DMACA (les résultats exprimés en μg équivalent catéchine par gramme de matière sèche et le caractère est noté estdma).

- Caractères mécaniques :

La nervosité du bois représente l'ensemble des propriétés du bois qui entraînent des déformations pendant la transformation, en l'absence d'efforts extérieurs (Baillères, 1994 [2]). Lors de la croissance, les nouvelles cellules, élastiques, sont soumises à

des contraintes du fait de leur accolement à des cellules plus anciennes et rigidifiées. Ainsi, des contraintes appelées "contraintes de croissance" apparaissent. Lors de l'abattage des arbres, ces contraintes se libèrent.

Pour étudier la nervosité, deux caractères mécaniques ont été utilisés : les déformations résiduelles longitudinales de maturation (DRLM) et l'indice d'éclatement (IE). Les DRLM ont été mesurées, 1 semaine après abattage, par la méthode du trou unique grâce à un appareil développé au Cirad (Mariaux, 1982 [35]). L'indice d'éclatement est défini par la somme des longueurs totales des fentes observées sur la coupe transversale et sur la périphérie de la section d'un billon rapporté au diamètre de la section. L'indice d'éclatement a été calculé à différentes dates après abattage : 1 jour (IE1j), 1 semaine (IE1Sem) et 2 mois (IE2Mois).

- Caractères physiques :

Les caractères physiques correspondent aux mesures de densité du bois obtenue par une mesure indirecte (pilodyn) et de rigidité du bois par l'estimation du module d'élasticité spécifique. La pénétration du pilodyn dans le bois permet d'estimer indirectement la densité du bois en périphérie du tronc. En effet, plus le bois est dense, moins l'aiguille s'enfonce. Cette mesure a été réalisée à 51 mois après écorçage. Le module d'élasticité est défini par le rapport entre la contrainte imposée à une pièce de bois et la déformation qui en résulte. Le module d'élasticité spécifique correspond au rapport entre le module d'élasticité et la masse volumique, c'est à dire que cette mesure s'affranchit de la porosité (Gibson *et al.*, 1999 [20]). La mesure utilisée est notée ESmoy30 et représente la valeur moyenne du module d'élasticité spécifique à 30 mois.

- Caractères anatomiques :

Les caractéristiques anatomiques du bois ont été étudiées au travers des caractéristiques morphologiques des fibres de cellulose. Les dimensions de ces fibres conditionnent la qualité du papier (déchirement, souplesse et opacité). Différentes mesures ont été effectuées : masse des fibres en μ /mm (masselinéiq), longueur moyenne des fibres en mm après élimination des fibres cassées, largeur des fibres en μ m (largeur) et l'élancement des fibres qui est le rapport entre la longueur des fibres et leur diamètre, exprimé en pourcentage(souplesse).

2.7.2 Statistiques descriptives des variables

Pour chaque caractère, la moyenne et l'écart-type ont été calculés sur l'ensemble de la descendance (200 individus). Pour pouvoir comparer la variation des caractères ayant des unités de mesure différentes, le coefficient de variation (écart-type / moyenne) a été calculé.

Le coefficient de corrélation de Pearson et la probabilité associée ont été calculés pour chaque couple de variables étudiées grâce au logiciel Q-gene version 3.0 (Nelson, 1997 [40]). Ce coefficient traduit la relation linéaire entre deux variables, ici deux caractères quantitatifs.

2.8 Détection de QTL

La détection d'un QTL correspond à la mise en évidence d'une liaison statistique entre une variation génotypique (ségrégation d'un marqueur dans une descendance) et une variation phénotypique (variation d'un caractère quantitatif dans cette même descendance). Il existe plusieurs méthodes de détection : des méthodes telle que l'analyse de variance qui permet une détection marqueur par marqueur, et des méthodes dite "d'Interval Mapping", utilisées lorsqu'une carte génétique est disponible permettant de préciser la localisation et l'effet des QTL.

2.8.1 Détection marqueur par marqueur (ANOVA)

Deux modèles d'analyse de variance ont été utilisés, un modèle à un facteur et un modèle à deux facteurs avec interaction, selon que les EST étaient polymorphes pour un seul ou les deux parents du croisement.

- Analyse de variance à 1 facteur :

Pour une EST polymorphe chez un géniteur de génotype "12", les descendants hybrides peuvent être regroupés en deux classes alléliques : une possédant l'allèle "1" au marqueur et l'autre possédant l'allèle "2". Les moyennes de ces 2 classes sont comparées pour le caractère étudié, grâce au modèle de variance à un facteur :

$$Y_{ij} = \mu + M_i + \varepsilon_{ij}$$

Avec Y_{ij} la valeur du caractère observé pour l'individu j de génotype i (1 ou 2) au marqueur, μ la moyenne du caractère pour la descendance, M_i l'effet moyen du génotype i au marqueur et ε_{ij} le terme résiduel.

L'effet du marqueur étudié est testé grâce à un test F, au seuil de 1%. Le coefficient de détermination ($R^2 = \text{SCE facteur} / \text{SCE totale}$) a été utilisé pour estimer la part de variation phénotypique expliquée par le QTL. Ces analyses ont été effectuées grâce au Logiciel R (disponible sur <http://cran.r-project.org/>).

- Analyse de variance à 2 facteurs :

Ce test est réalisé pour les EST polymorphes chez les deux parents. Il permet d'analyser les effets des allèles maternels et paternels ainsi que d'éventuelles interactions entre les allèles parentaux dans la descendance. Le modèle utilisé se décompose comme suit :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \varepsilon_{ijk}$$

Avec Y_{ijk} la valeur du caractère observé pour l'individu k portant l'allèle maternel i (1 ou 2) et l'allèle paternel j (1 ou 2), μ la moyenne du caractère pour la descendance, α_i l'effet moyen de l'allèle maternel i au locus considéré, β_j effet moyen de l'allèle paternel j au locus considéré, δ_{ij} l'effet d'interaction entre les deux allèles parentaux et ε_{ijk} le terme résiduel.

2.8.2 Détection par Interval Mapping (IM)

La détection de QTL par IM est possible car nous possédons les cartes génétiques pour les deux géniteurs du croisement. L'IM part du postulat qu'il existe au maximum un QTL entre deux marqueurs. La présence du QTL est testée en chaque point de la carte (chaque centimorgan) par l'estimation d'un LOD score, qui correspond au logarithme décimal du rapport de vraisemblance :

$$LOD = \log_{10} \left[\frac{V1}{V2} \right]$$

Avec : V1 la valeur de la fonction de vraisemblance sous l'hypothèse de présence d'un QTL V2 la valeur de la fonction de vraisemblance sous l'hypothèse d'absence de QTL.

Les valeurs V1 et V2 tiennent normalement compte des effets additifs et de dominance des QTL, cependant, dans le cadre de cette étude, ces effets ne peuvent être calculés indépendamment. L'évolution des LOD est suivie tout le long des groupes de liaison, et lorsqu'il dépasse un seuil, fixé ici à 2, on déclare l'existence d'un QTL. Le maximum de LOD observé correspond à la position la plus probable du QTL. L'intervalle de confiance sur la position du QTL correspond à une diminution de

LOD de 1 unité par rapport au LOD maximum. La détection de QTL est réalisée indépendamment pour chaque parent du croisement avec Le logiciel Qgene version 3.0 (Nelson, 1997 [40]).

2.8.3 Tests de permutations

Nous avons utilisé des tests de permutations afin de valider l'existence des QTL détectés par IM. Le test de permutation fournit par QGENE est basé sur les travaux de Churchill et Doerge (1994 [13]) et consiste en une analyse marqueur par marqueur. Pour chaque marqueur, les génotypes sont redistribués au hasard sur la totalité de l'échantillon. Cette opération est répétée 10000 fois (10000 permutations) et à chaque répétition, un test F est réalisé. Les permutations créent de façon empirique une distribution de F pour l'hypothèse nulle H_0 : "absence de QTL". Si la valeur du F observée est supérieure au F seuil à 1%, l'hypothèse nulle est rejetée avec un risque d'erreur $\alpha = 0,01$. On déclare donc la présence d'un QTL avec ce risque.

Tableau 2 : Conditions PCR pour les 24 EST qui ont pu être amplifiées et séquences des amorces spécifiques utilisées. U: amorce upper; L: amorce lower. Un * désigne qu'un intron a été amplifié.

EST Id.	Séquences des amorces (5'→3') définies sur les EST d' <i>E.gunnii</i>	Taille attendue en pb	Taille observée en pb	T(°C) d'hybridation
6A01	U: AAA TCG AGC TGG TGC CTG AC L: AGT AGC TGG TGG GTT CCT GT	375	350-400	63
6A05	U: GCG GGC CGA GGG TCA CTG L: GTT GAT AAT GTG GAT GCC	330	300-350	58
6A07	U: AGA CGA AGC AGG AGG AGG TGA T L: TGT ATT AGT TTT GGC CTG TTT C	232	200-250	65
6A08	U: AAG CAC ACA ACA ACA CAA ACA L: ACT TCC TCC CTT TTC AGC CCC	375	350-400	58
6A11	U: GAT AGT GAA AAG TGA GGC AAT GG L: CTA CTG CTA CGA CAC TCT GCG GT	332	300-350	60
6A12	U: GGG AAC GAG GAC ATA CC L: AAG TGT CTG TGG CTC CA	144	150	56
6B08	U: GGA GAA GAA AGC GGA GAC L: TAG CGA GGG AGA TGA TGA	226	200-250	60
6B07	U: GGG TCT TTT GTT TCG TTC G L: TCC CAC AAG AAT CCA AAC C	127	100-150	52
6C08	U: CAA GCT GCT GAG GAT AGT C L: GAA AAC ACT CTG CAA AAT G	120	100-150	56
6D01	U: TGT TGG CTG GCA TTT CCT TCC L: CGG TAG GTG AAT GCC CAG AGG	295	300	68
6D02	U: CAG GAT GCG AGG GTG AGA CAG L: AGT CTT CCT TCT CTC GGC TTG	192	200	60
6D03	U: TCA GGG AAG CAA ATA ACT L: GTT TTG CTA TTT TGA GGC	184	150-200	50
6D07	U: GAA GGC TTA GAG ACG GGT TTT L: AAA ACC AGT GCC TCA ACC AGT	147	550*	60
6D12	U: TTG CAG TTC AGC CCC CTC C L: ATG CGA AGC GAC TCC CCT C	196	700-800*	54
6F06	U: CAG GGG ATG TTC AAG ACG G L: AGG ATC AAC GGA AGC CAG C	240	200-250	60
6H12	U: GCG GTG AGG AAC GAC GAG GAG L: TGG CTT TGG ACG GAG ACT TGG	145	150	70
7A01	U: GTG GAA TGT TCA TCT CGC L: GTG AAA TAT GGT GCT GGC	352	300-400	56
7A07	U: GCA AAG GCA ATA GGG AGC AGC L: AGG AAG AAG GGC AAG TGG AGG	333	400-500*	60
7A08	U: TCA AAA GGA CAA GAA CAC G L: ACA GGC AGT AAC TTC CGA C	144	150	56
7A09	U: GAC TGT CCC TAC CCT TGC L: GCT CTC TTA TCC CCT TTG	143	100-150	56
9G11	U: GGA GGA GGA GGG CGA GAC G L: GCG GCG GAG GAG AAT GGA C	400	600-700*	67
11H04	U: ATG CGT TAT TAC ACA AGG L: GAT TTC AAG TGG GTC AAC	521	500-600	59
12A02	U: TAC AAC AAC GGC AAA GTG L: TCA CCA TCC ATC TCC ATC	245	500*	60
12A09	U: ACC TGT CCC TAT GTT TGT L: AGC CGT GAA TCT TTA GTG	245	300-400*	49

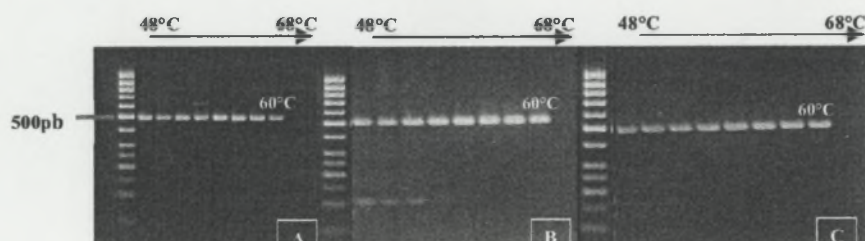


Figure 4 : Résultats d'électrophorèse sur gel d'agarose à 2% de produits d'amplification de l'EST 12A02, pour un gradient de températures allant de 48°C à 68°C. A: profils obtenus pour *E. urophylla*; B: profils obtenus pour *E. grandis*; C: profils obtenus pour *E. globulus*. Un marqueur de taille sépare A, B et C.

3 RESULTATS

3.1 Cartographie génétique des EST

3.1.1 Amplification des EST

L'ensemble des couples d'amorces spécifiques a été testé sur trois génotypes d'*Eucalyptus*: *E. globulus* (individu CN32 de RAIZ), ainsi que les 2 parents du croisement interspécifique étudié *E. urophylla* et *E. grandis*. Lors de cette phase d'optimisation des conditions PCR, seule la température d'hybridation a été testée.

Sur les 34 EST testées, 24 couples d'amorces ont donné un produit d'amplification unique (une seule bande sur gel d'agarose) chez les 2 géniteurs *E. urophylla* et *E. grandis*, ainsi que chez *E. globulus* (tableau 2). Parmi ces 24 produits d'amplification uniques, 18 ont une taille observée identique à celle attendue avec la séquence d'*E. gunnii*. Les 6 autres ont une taille supérieure à la taille attendue laissant supposé qu'un intron ait été amplifié. La figure 4 montre les profils obtenus sur ces 3 génotypes pour l'EST 12A02, pour un gradient de température allant de 48 à 68,5°C. Pour cette EST, la température d'hybridation optimale est de 60°C, température à laquelle un fragment d'environ 500 pb (taille attendue de 245 pb) est amplifié pour les trois génotypes testés. On suppose donc qu'un intron d'environ 250 pb a été amplifié pour l'EST 12A02. Lorsqu'un seul fragment de taille similaire est obtenu chez les 3 génotypes d'espèces différentes, le produit d'amplification est supposé correspondre à l'EST ciblée. Seule une étape de séquençage, non réalisée au cours de mon stage, pourrait confirmer cette hypothèse.

Les 10 derniers couples d'amorces n'ont pas permis d'obtenir un produit d'amplification satisfaisant pour les trois espèces (pas d'amplification ou profil multibande) ou seulement pour le génotype *E. globulus* (12% des amorces testées). Ces différences entre espèces sont sans doute dues au fait qu'*E. globulus* et *E. gunnii* appartiennent à la même section (*Maidenaria*) alors qu'*E. urophylla* et *E. grandis* appartiennent à la section *Transversaria*.

3.1.2 Détection du polymorphisme par SSCP

Une première étape d'optimisation des conditions SSCP pour la mise en évidence de polymorphisme a été réalisée avec les deux géniteurs (*E. urophylla* et *E. grandis*) et six descendants hybrides. Cette étape préliminaire a lieu pour chaque EST. Une

Tableau 3 : Conditions utilisées en SSCP pour les différentes EST et polymorphisme observé pour les deux géniteurs. P indique que l'EST est polymorphe pour le géniteur considéré et M que l'EST est monomorphe.

EST id.	Taille des fragments amplifiés en pb	Température de migration (T°C)	Concentration en acrylamide	Voltage	Durée de migration	<i>E. grandis</i> : 9-21	<i>E. urophylla</i> : 14-144
6A07	200-250	4°C	12%	350	7h	P	P
6A11	300-350	15°C	12%	250	17h	P	M
6A12	150	15°C	12%	300	5h30	P	M
6B08	200-250	15°C	12%	350	7h30	P	M
6D01	300	15°C	12%	250	15h	P	P
6D03	150-200	15°C	12%	200	14h	M	P
6D07	550	15°C	12%	250	23h	M	P
6D12	700-800	15°C	12%	350	22h	M	P
7A01	300-400	15°C	12%	200	16h30	M	P
7A07	400-500	15°C	12%	300	20h	P	P
7A08	150	15°C	12%	250	6h	P	M
7A09	100-150	15°C	12%	300	6h	P	P
11H04	500-600	4°C	12%	350	21h	P	M
12A02	500	15°C	12%	350	21h	P	M

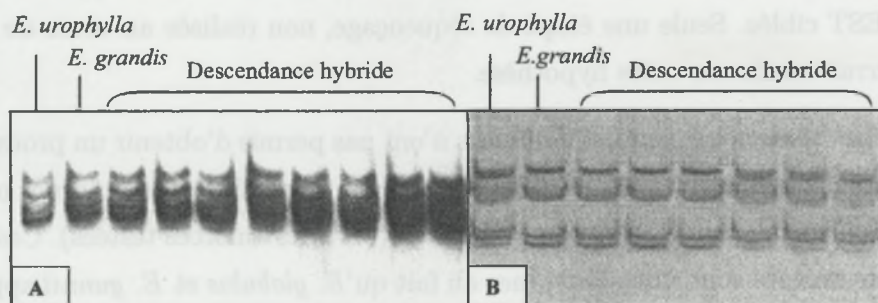


Figure 5 : Comparaison de 2 profils SSCP obtenus pour la même EST 6D03, avant (A) et après (B) optimisation des conditions de migration. Conditions pour A : gels à 8% d'acrylamide, migration 5h30 à 20°C et 200V. Conditions pour B : gels à 12% d'acrylamide, migration 14h à 15°C et 200V.

fois ces conditions optimales déterminées, la mise en évidence du polymorphisme est effectuée pour les deux deux géniteurs et les 92 hybrides.

- Optimisation des conditions de migration :

La SSCP est une méthode qui nécessite une phase d'optimisation car elle est sensible à différents paramètres qui influencent la révélation du polymorphisme. Ainsi, différentes conditions ont été testées : température de migration, concentration du gel en acrylamide, présence ou absence de glycérol dans le gel. Suivant la taille des fragments amplifiés, la durée de migration et le voltage ont aussi été adaptés.

Dans notre étude, la présence de glycérol dans les gels (5%) n'a pas facilité l'interprétation des résultats. Une concentration de 12% d'acrylamide a permis d'obtenir des profils plus lisibles et avec une meilleure séparation des bandes. Cette concentration a été utilisée pour toutes les EST. Il a été montré que la température de migration était un facteur important pour la résolution des gels (Fukuoka *et al.*, 1994 [19]). Les électrophorèses ont donc été réalisées à 3 températures différentes : 4, 15 et 20°C et on observe que la diminution de la température de migration permet une lecture plus facile des profils. De ce fait, la température de migration a été choisie au cas par cas pour chaque EST. A titre d'exemple, deux profils obtenus pour l'EST 6D03, avant et après optimisation des conditions SSCP, sont présentés dans la figure 5. Les premières conditions testées (8% d'acrylamide, migration à 200V, 20°C et 5h30) ne permettaient pas de détecter du polymorphisme. Pour les conditions finales (12% d'acrylamide, migration à 200V, 15°C et 14h) on met en évidence des bandes issues du géniteur *E. urophylla* qui ségrègent dans la descendance. Les différentes conditions utilisées pour chaque EST sont détaillées dans le tableau 3.

L'optimisation des conditions SSCP pour les 24 EST a permis d'en révéler 4 polymorphes chez les deux parents, 7 polymorphes uniquement chez *E. grandis* et 3 chez *E. urophylla*. Enfin, 10 EST semblent monomorphes quelles que soient les conditions testées. Pour ces dernières, d'autres tests pourraient être réalisés en optimisant d'autres paramètres. Le fragment amplifié peut également ne pas présenter de polymorphisme. Dans ce cas, il faudrait définir une autre paire d'amorces pour ces EST, de manière à augmenter les chances de détecter du polymorphisme. Enfin, l'EST étudiée peut être monomorphe dans le croisement étudié et ne pas être exploitable pour les études de cartographie génétique et de détection de QTL.

Tableau 4 : Interprétation des profils SSCP obtenus pour les 14 EST polymorphes et test de la ségrégation mendélienne par un test du χ^2 à 5%. (a) Valeur de χ^2 observé pour chaque EST. Pour $\alpha = 5\%$, la valeur seuil du χ^2 est de 3,84 pour 1 degré de liberté et de 7,81 pour 3 degrés de liberté. (b) Un - indique un allèle nul.

EST	Génotypes parentaux <i>E. grandis</i> / <i>E. urophylla</i>	Génotypes des hybrides	Effectifs observés	Effectifs théoriques	$\chi^2(a)$	ddl
6A07	(12;34) / (56;78)	(12;56) / (12;78) / (34;56) / (34;78)	20 / 22 / 15 / 29	21.50 / 21.50 / 21.50 / 21.50	1.67	3
6A11	(12;34) / (12;12)	(12;34) / (12;12)	47 / 37	42 / 42	1.19	1
6A12	(12;34) / (56;56)	(12;56) / (34;56)	36 / 49	42.5 / 42.5	1.98	1
6B08	(12;-) ^b / (34;34)	(12;34) / (-;34)	37 / 52	44.5 / 44.5	2.53	1
6D01	(12;-) / (34;56)	(12;34) / (12;56) / (-;34) / (-;56)	23 / 24 / 22 / 20	22.25 / 22.25 / 22.25 / 22.25	0.005	3
6D03	(12;12) / (12;34)	(12;12) / (12;34)	41 / 50	45.5 / 45.5	0.89	1
6D07	(12;34) / (12;12)	(12;34) / (12;12)	25 / 39	32 / 32	3	1
6D12	(12;12) / (34;56)	(12;34) / (12;56)	37 / 44	40.5 / 40.5	0.6	1
7A01	(12;34) / (12;12)	(12;34) / (12;12)	41 / 37	39 / 39	0.2	1
7A07	(12;34) / (56;78)	(12;56) / (12;78) / (34;56) / (34;78)	26 / 25 / 12 / 20	20.75 / 20.75 / 20.75 / 20.75	2.47	3
7A08	(12;34) / (12;12)	(12;34) / (12;12)	48 / 42	45 / 45	0.39	1
7A09	(12;34) / (56;78)	(12;56) / (12;78) / (34;56) / (34;78)	26 / 23 / 27 / 14	22.5 / 22.5 / 22.5 / 22.5	1.66	3
11H04	(12;34) / (12;12)	(12;34) / (12;12)	41 / 44	42.5 / 42.5	0.1	1
12A02	(12;34) / (12;12)	(12;34) / (12;12)	42 / 46	44 / 44	0.18	1

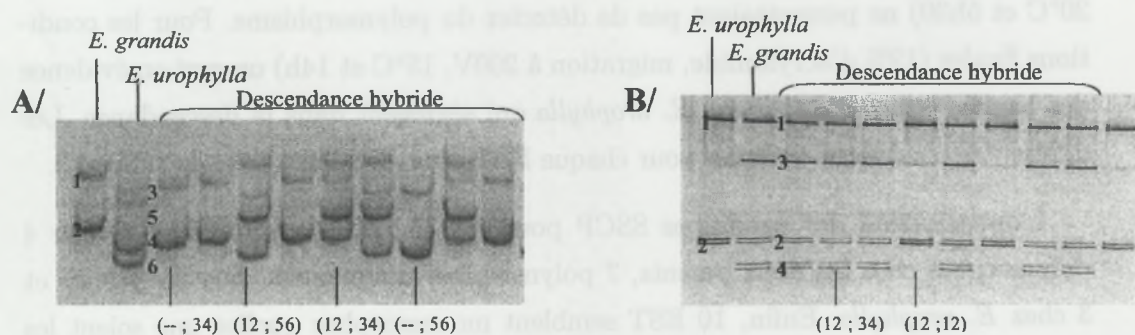


Figure 6 : Révélation du polymorphisme pour deux EST étudiés. A/ EST 6D01 : Le parent mâle *E. grandis* est (12;-) et le parent femelle *E. urophylla* est (34;56). Les descendants sont : (12;34), (12;56), (-;34) et (-;56). B/ EST 7A01: Le parent mâle *E. grandis* est (12;34) et le parent femelle *E. urophylla* est (12;-). Les descendants ont les génotypes (12;34) et (12;12).

- Ségrégations observées pour les EST polymorphes :

Le profil SSCP obtenu pour l'EST 6D01 montre 2 bandes chez *E. grandis* (12) et 4 bandes chez *E. urophylla* (3456) (figure 6). Dans la descendance, 4 classes d'individus sont observées correspondant à la ségrégation des allèles parentaux. En effet, on observe les classes alléliques (34), (56), (12;34) et (12;56). L'observation de quatre classes de descendants signifie que les deux parents sont hétérozygotes au locus considéré. Les deux allèles d'*E. urophylla* sont représentés par les bandes 3 et 4 pour l'un et 5 et 6 pour l'autre (34;56). Pour *E. grandis* qui ne présente que 2 bandes (12), un allèle nul, c'est à dire qui n'est pas amplifié, n'est pas représenté mais il ne gêne pas l'observation d'une ségrégation des allèles d'*E. grandis* dans la descendance. Le génotype d'*E. grandis* est noté (12;-).

Pour des génotypes diploïdes (cas des eucalyptus), en théorie 4 bandes au maximum doivent être observées en SSCP chez un individu hétérozygote. Néanmoins, pour l'EST 7A01, le parent *E. grandis* présente 6 bandes différentes comme indiqué sur la figure 6. Dans la descendance, deux classes d'individus sont observées correspondant à la ségrégation des allèles de ce géniteur. Pour cette EST, *E. grandis* semble bien hétérozygote avec plusieurs conformations simple brin possibles pour une même forme allélique. Ce résultat peut être expliqué par une instabilité de certaines conformations simple brin qui, dans les conditions SSCP utilisées, génère plusieurs niveaux de bandes pour un même allèle.

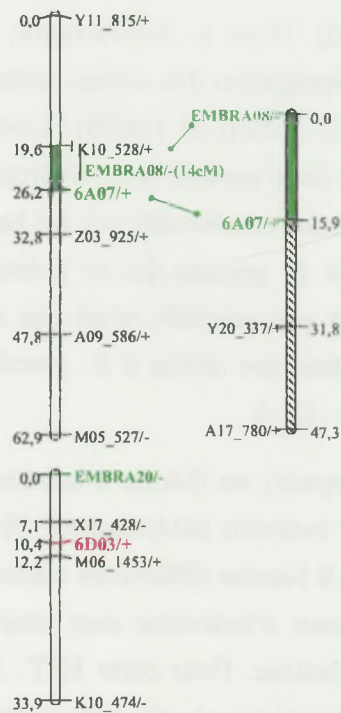
Pour les 14 EST polymorphes chez au moins un des parents du croisement, plusieurs types de profils ont été observés : 2 bandes, 3 bandes, 4 bandes ou encore un profil multi-bandes. Les ségrégations observées des allèles parentaux dans la descendance sont en adéquation avec une ségrégation mendélienne au seuil de 5% (tableau 4).

3.1.3 Localisation des EST sur les cartes génétiques

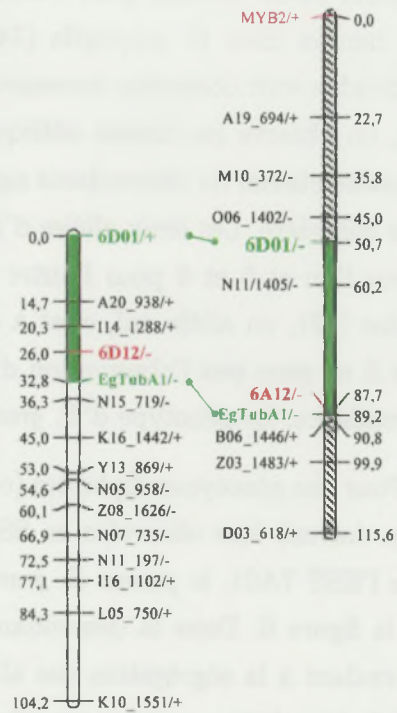
Les EST polymorphes ont été cartographiées sur les cartes génétiques d'*E. grandis* et d'*E. urophylla* déjà établies par Gion (2001 [21]) à partir des marqueurs dominants de type RAPD, de locus microsatellites et de gènes de fonction connue et des 201 hybrides du croisement élite.

Les liaisons génétiques entre les EST étudiées et les marqueurs jalons des cartes sont très significatives avec des valeurs de LOD score comprises entre 8 et 22, sauf pour l'EST 6A11 qui est plus faiblement liée (LOD de 3,5) et l'EST 7A09 qui n'est

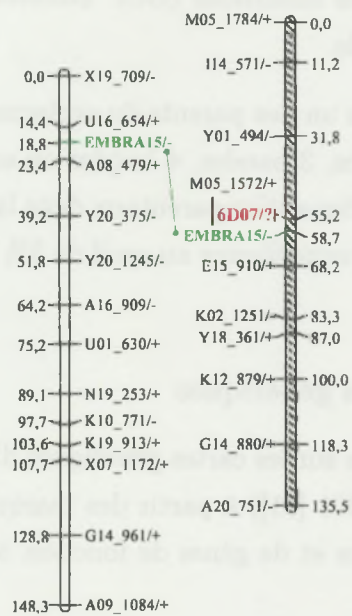
Groupe de Liaison 1



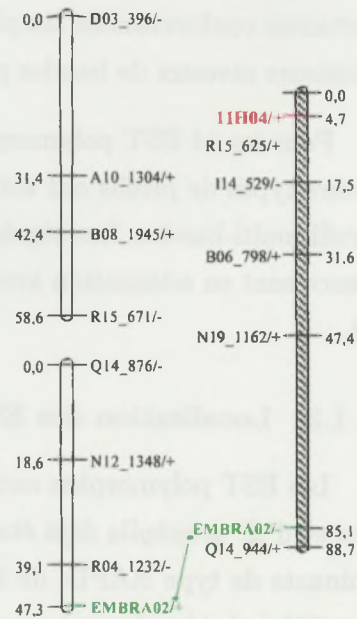
Groupe de Liaison 2



Groupe de Liaison 3



Groupe de Liaison 4



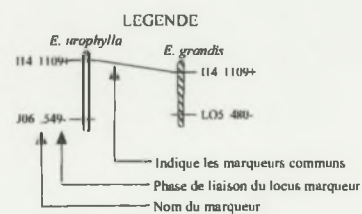
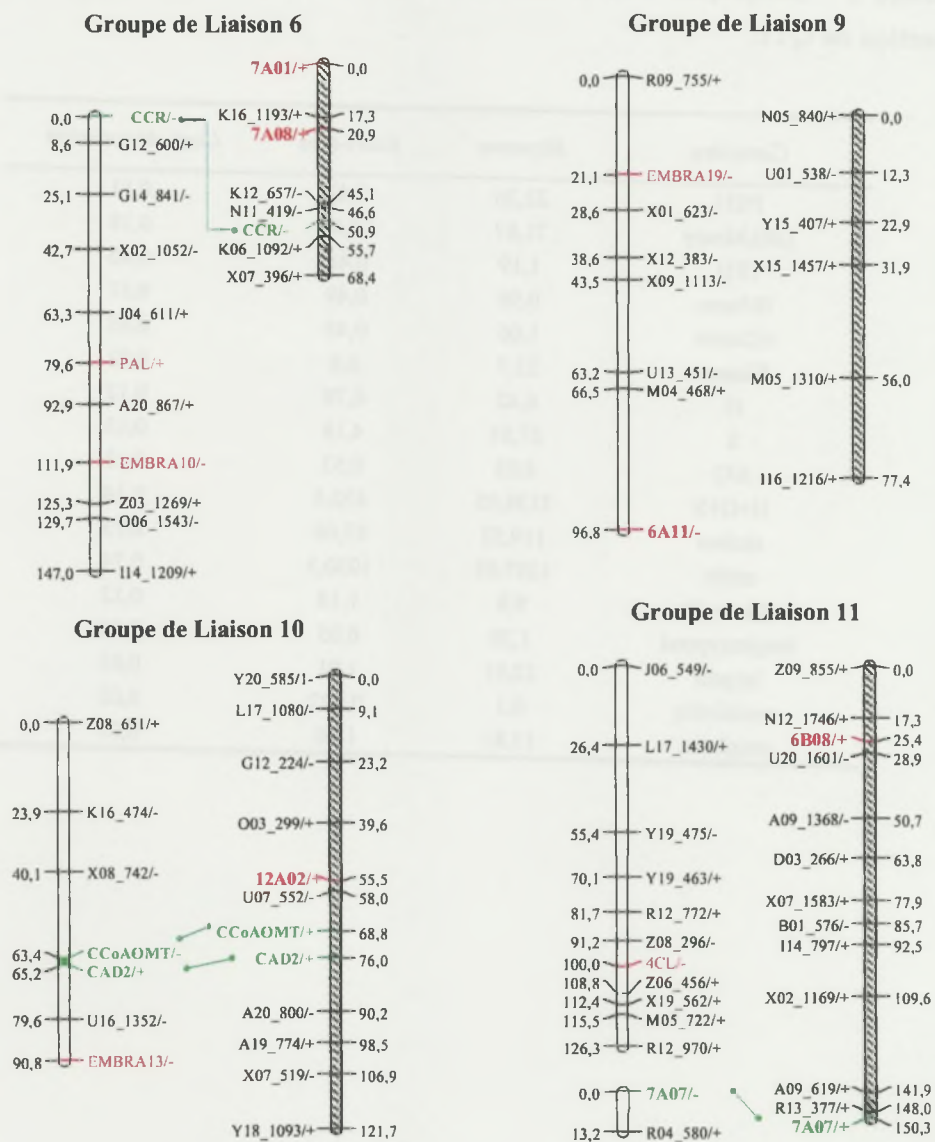


Figure 7 : Localisation des EST polymorphes sur les cartes génétiques d'*E. urophylla* et d'*E. grandis*.

Tableau 5 : Statistiques descriptives des caractères de qualité du bois étudiés pour la détection de QTL.

<i>Caractère</i>	<i>Moyenne</i>	<i>Ecart-type</i>	<i>Coef. de variation</i>
PilS1	22,26	2,42	0,11
DRLMmoy	71,67	12,68	0,18
IE1j	1,19	0,48	0,40
IE1sem	0,96	0,49	0,51
IE2mois	1,06	0,48	0,45
Klason	23,7	0,8	0,03
G	6,42	0,78	0,12
S	27,55	4,18	0,15
S/G	4,03	0,53	0,13
H+G+S	3138,05	452,5	0,14
estdma	119,52	87,09	0,73
estfol	1397,03	1030,3	0,74
ESmoy30	9,9	1,14	0,12
longmoypond	1,28	0,05	0,04
largeur	32,91	1,01	0,03
masslinéiq	0,1	0,002	0,02
souplesse	15,81	1,06	0,07

liée à aucun des marqueurs jalons d'*E. grandis* et d'*E. urophylla*. Deux hypothèses peuvent être formulées pour expliquer la non liaison de cette EST avec les marqueurs des cartes génétiques parentales: (i) une mauvaise interprétation des ségrégations alléliques, (ii) le niveau de saturation des cartes génétiques n'est pas suffisant. Pour la majorité des EST, la localisation sur les groupes de liaisons existants a été possible chez *E. grandis* et *E. urophylla* (figure 7).

- Cartographie des EST chez *E. grandis* :

Pour ce géniteur, sur les 11 EST polymorphes, 10 ont pu être positionnées sur les différents groupes de liaison. L'EST 6A07 a été localisée sur le groupe de liaison 1 (GL1), 6A12 et 6D01 ont été cartographiées sur GL2, 6D07 sur GL3, 11H04 sur GL4, 7A01 et 7A08 sur GL6, 12A02 sur GL10, enfin, 6B08 et 7A07 sur le GL11.

- Cartographie des EST chez *E. urophylla* :

Chez *E. urophylla*, sur les 7 EST polymorphes, 5 ont été localisées sur des groupes de liaisons existants. Les EST 6A07 et 6D03 sont positionnées sur GL1 (a et b). Les EST 6D01 et 6D12 sont liées et localisées sur le GL2 et 6A11 est positionné sur le GL9. L'EST 7A07 présente une liaison significative avec le marqueur RAPD R04580/+ (LOD de 11,37) qui n'est pas positionné sur un des 11 groupes de liaison de ce géniteur.

3.2 Analyse des colocalisations EST/QTL

3.2.1 Variabilité des caractères quantitatifs étudiés

La mise en évidence de QTL pour un caractère nécessite que le caractère étudié présente une variabilité dans la descendance. Les caractères de qualité du bois étudiés (chimiques, mécaniques, anatomiques et physiques) présentent tous une variabilité significative au sein de la descendance étudiée. Les statistiques descriptives des différentes variables sont données dans le tableau 5. Pour le caractère d'indice d'éclatement les valeurs obtenues aux différents age sont homogènes, avec une moyenne aux alentours de 1 et un écart type important (0,48 et 0,49). Les coefficients de variation sont importants, compris entre 0,40 et 0,51. Les teneurs en phénols et en flavanols sont les variables présentant le plus de variabilité dans la descendance étudiée, avec respectivement des coefficients de variation de 0,74 et 0,73. Il existe donc une forte hétérogénéité entre les individus au sein du croisement étudié. On observe également une différence importante de la teneur en monomères S ($27,5 \mu\text{mol/g}$) par rapport

Tableau 6 : Corrélations entre les différents caractères de qualité du bois étudiés. Les corrélations significatives au seuil de 1% sont indiquées en gras.

Trait	Pil51	DRLM	IEtot1J	IEtot1S	IEtot2M	Klason	longmoy	largeur	massliné	souplesse	G	S	H+G+S	S/G	ESmoy30	estfol
DRLM	-0.034															
IEtot1J	-0.135	0.336														
IEtot1S	-0.118	0.411	0.738													
IEtot2M	-0.118	0.364	0.380	0.604												
Klason	-0.083	-0.112	0.086	-0.058	-0.098											
longmoy	-0.102	0.215	0.146	0.202	0.119	-0.140										
largeur	-0.006	-0.036	-0.081	-0.096	-0.126	0.284	-0.458									
massliné	-0.109	0.124	0.091	0.167	0.090	-0.120	0.729	-0.311								
souplesse	-0.012	-0.045	-0.050	-0.087	-0.078	0.302	-0.611	0.954	-0.500							
G	-0.007	0.008	-0.057	-0.050	0.094	0.133	0.081	0.058	-0.016	0.034						
S	0.273	-0.075	-0.184	-0.227	0.013	-0.200	-0.007	0.004	-0.066	-0.003	0.549					
H+G+S	0.244	-0.023	-0.173	-0.174	0.051	-0.379	0.052	-0.067	-0.019	-0.080	0.600	0.969				
S/G	0.322	-0.099	-0.158	-0.222	-0.082	-0.337	-0.081	-0.048	-0.060	-0.034	-0.322	0.610	0.523			
ESmoy30	-0.442	0.256	0.209	0.266	0.245	-0.167	0.178	-0.272	0.091	-0.192	-0.107	-0.224	-0.155	-0.170		
estfol	0.056	-0.005	0.025	0.024	0.065	0.142	-0.198	0.223	-0.155	0.236	-0.034	0.129	0.037	0.180	-0.177	
estdma	-0.051	-0.003	0.068	-0.058	-0.022	0.132	-0.098	0.172	-0.114	0.184	-0.096	0.093	0.002	0.191	-0.075	0.636

Tableau 7 : Résultats de l'analyse de variance à 2 facteurs. Les probabilités associées aux effets des allèles maternels (pv/alleleM), paternels (pv/alleleP) et à leurs interactions (pv/inter) sont fournies, ainsi que la variation phénotypique expliquée par le QTL (R^2). Les QTL détectés au seuil de 1% sont indiqués en gras.

<i>Caractère</i>	<i>EST</i>	<i>pv/AllelM</i>	<i>pv/AllelP</i>	<i>pv/inter</i>	<i>r²</i>
G	6A07	0,00829	0,79124	0,03606	0,13611
G	7A09	0,12969	0,00495	0,34431	0,12667
S	7A09	0,45010	0,00320	0,16609	0,12823
estdma	7A07	0,18067	0,00297	0,45431	0,13788
ESmoy30	6D01	0,00191	0,02873	0,63990	0,16040
Pil51 sous	6D01	0,00059	0,02872	0,46477	0,17661
longmoypond	6D01	0,00041	0,11194	0,99650	0,17195
largeur	6D01	0,04175	0,00920	0,44276	0,13348
masslineiq	6D01	0,00273	0,19924	0,39237	0,13332

aux monomères G ($6,42 \mu\text{mol/g}$). Cette différence est caractéristique des feuillus, donc des eucalyptus, alors que les bois de conifères sont enrichis en unités G.

Les coefficients de corrélation de Pearson entre les différents caractères de qualité du bois sont donnés dans le tableau 6. Les variables appartenant aux mêmes types de caractères sont corrélés entre eux. Ainsi, la DRLM est corrélée avec les indices d'éclatement, de même que les différentes caractéristiques des fibres, les différents caractères de teneur en lignines et les caractères physiques. Aucune corrélation n'est observée entre le taux de lignines et la DRLM, or, Baillères (1994 [2]) avait mis en évidence une corrélation négative entre ces 2 caractères. La longueur moyenne pondérée des fibres de bois est corrélée significativement avec les 2 caractères qui traduisent la nervosité du bois. Une corrélation significative est observée pour le pilodyn, caractère de densité du bois et la teneur en monomère S de la lignine, le rapport S/G, ainsi que la somme des différents monomères de lignine, alors qu'aucune corrélation n'est mise en évidence avec la teneur en G. Cette corrélation indique qu'un bois nerveux se caractérise par une teneur en lignine en S faible. Enfin, le module spécifique d'élasticité est corrélé significativement avec beaucoup d'autres caractères. Il existe notamment une corrélation négative avec le pilodyn, indiquant que plus le bois est dur, plus l'élasticité est importante.

3.2.2 Analyse QTL marqueur par marqueur

- Analyse de variance à 1 facteur :

Cette analyse a été réalisée pour les EST polymorphes chez un seul des parents du croisement. Chez *E. grandis*, au seuil de 1%, aucun QTL n'a pu être mis en évidence pour les 7 EST testées. Pour *E. urophylla*, l'EST 6D12 présente un effet sur le caractère pilodyn à 51 mois. La probabilité associée est de 0,005 et le coefficient de détermination R^2 , qui représente la proportion de la variance phénotypique expliquée par le QTL est de 9,5%.

- Analyse de variance à 2 facteurs :

Le modèle statistique permet d'analyser conjointement l'effet des allèles maternels et paternels, ainsi que leur éventuelle interaction. Pour cette analyse, les EST 6A07, 7A07, 7A09 et 6D01 ont été utilisées car elles sont polymorphes chez les 2 parents.

Cette analyse a permis de détecter 9 QTL (tableau 7). Pour la teneur en monomères G, 2 QTL ont été détectés, un chez *E. urophylla* avec l'EST 6A07 et un chez *E. grandis* avec l'EST 7A09. Ces 2 QTL expliquent respectivement 13,6 et 12,6% de

Tableau 8 : Mise en évidence de QTL par Interval Mapping chez *E. grandis* (A) et chez *E. urophylla* (B) pour la qualité de bois : teneurs en lignines (Klason) et en monomères (G, S, H+G+S, S/G), teneurs en flavanols (estdma), la moyenne des déformations résiduelles longitudinales de maturation (DRLMmoy), pilodyn sous écorce à 51 mois (pil51), l'indice d'éclatement à 1 jour (IEtot1j), à 1 semaine (IEtot1sem), la longueur moyenne des fibres pondérées (longmoypond), le module d'élasticité spécifique à 30 mois (ESmoy30). Pour chaque QTL, le LOD max et la probabilité associée au test F (proba p), ainsi que le coefficient de détermination (R^2) sont donnés. La distance du QTL à l'EST la plus proche est également fournie (d(EST) en cM), de même que l'allèle favorable du marqueur le plus proche.

A/

Groupe de liaison	Caractère	Position (cM)	Marqueur flanquant	Allèle favorable	LOD max	Proba p	R^2	d(EST) en cM
1	longmoypond	0	E08_.000/1+	+	3,48	0,0001	0,0918	15,9
2	S	46	O06_1402/1-	+	2,20	0,0016	0,0522	4,7
	Pil51	45	O06_1402/1-	+	2,18	0,0016	0,0492	5,7
3	Pil51	55	6D07	-	2,96	0,0002	0,0662	<0,2
4	Klason	66	N19_1161/2+	-	2,69	0,0005	0,0634	61,3
	H+G+S	66	N19_1161/2+	-	2,03	0,0024	0,0483	61,3
	DRLMmoy	49	N19_1161/2+	+	2,70	0,0005	0,0609	44,3
	IEtot1sem	65	N19_1161/2+	+	2,08	0,0021	0,0487	60,3
6	S/G	68	X07_.396/2+	+	3,21	0,0001	0,0752	47,1
	S/G	41	K12_.657/1-	-	2,40	0,0010	0,0568	20,1
	Pil51	47	N11_.419/1-	-	3,42	0,0001	0,0761	26,1
10	Klason	87	A20_.800/1-	+	2,40	0,0010	0,0571	31,5
11	Klason	93	I14_.797/1+	-	2,02	0,0024	0,0480	57,3
	Klason	77	X07_1583/1+	-	2,07	0,0022	0,0494	51,6
	S/G	147	R13_.377/2+	+	3,12	0,0002	0,0732	3,3
	estdma	141	A09_.619/1+	+	3,86	>10-4	0,0926	9,3
	DRLMmoy	38	U20_1601/2-	-	2,52	0,0007	0,0569	12,6

B/

Groupe de liaison	Caractère	Position (cM)	Marqueur flanquant	Allèle favorable	LOD max	Proba p	R^2	d(EST) en cM
1a	G	9	Y11_.815/1+	+	3,74	>10-4	0,0846	17,2
	S/G	26	6A07_.000/1+	-	2,35	0,0011	0,0545	0,2
1b	longmoypond	23	M06_1453/1+	+	2,24	0,0014	0,0563	12,6
2	estdma	104	6D01_.000/1+	-	2,67	0,0005	0,0650	0,2
	ESmoy30	88	A20_.938/1+	-	4,07	>10-4	0,0935	16,2
	Pil51	104	6D01_.000/1+	+	2,69	0,0005	0,0604	0,2
	Pil51	79	6D12_.000/1-	-	2,59	0,0006	0,0582	0,8
	longmoypond	100	6D01_.000/1+	-	3,41	0,0001	0,0844	4,2
9	DRLMmoy	40	X12_.383/1-	-	2,18	0,0016	0,0494	56,8
	IEtot1j	44	X09_1113/1-	-	2,04	0,0023	0,0461	52,8

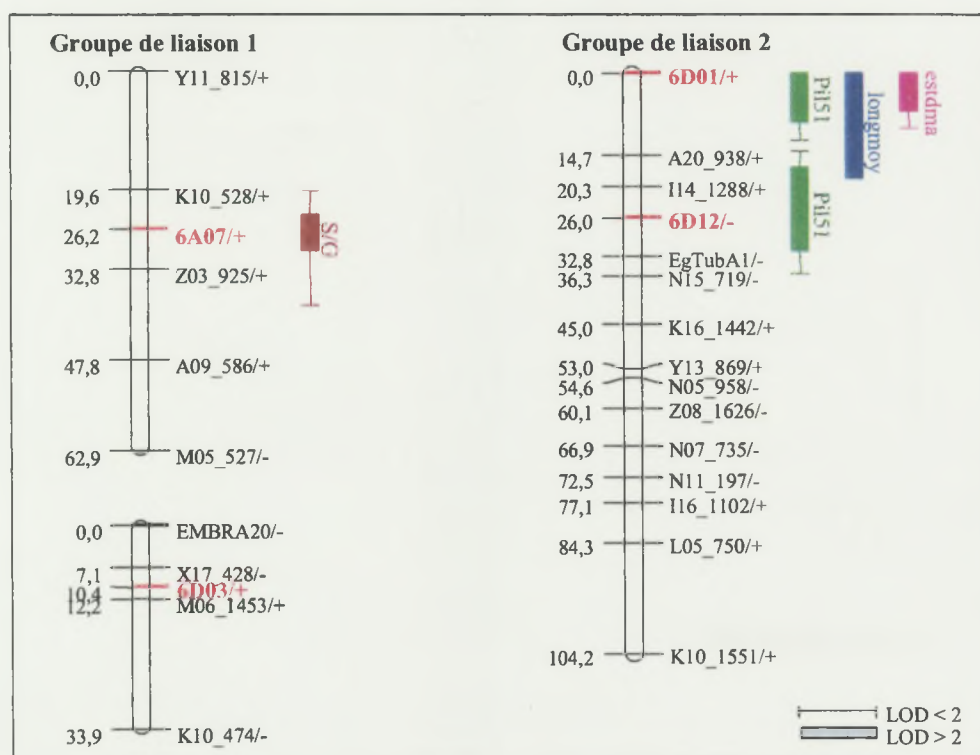


Figure 8 : Représentation de la position des QTL de qualité du bois détectés par Interval Mapping chez *E. urophylla*. Pil51: pilodyn sous écorce à 51 mois; longmoy: longueur moyenne pondérée des fibres de bois, estdma: teneur en flavanols et S/G rapport entre les monomères S et G.

la variance phénotypique. De plus, l'EST 7A09 colocalise avec un QTL de teneur en monomères S qui explique 12,8% de la variance phénotypique. L'EST 6D01 permet de détecter 5 QTL : le module d'élasticité spécifique à 30 mois, le pilodyn sous écorce à 51 mois, la longueur moyenne pondérée des fibres de bois et la masse linéique chez *E. urophylla*, la largeur des fibres chez *E. grandis*. Les variances phénotypiques expliquées par chacun des QTL pour ces différents caractères sont compris entre 13 et 20% (voir tableau 7). Lors de cette analyse, aucune interaction entre les allèles parentaux n'a été observée.

3.2.3 Analyse QTL par Interval Mapping (IM)

Afin de préciser l'effet de ces QTL et leur position sur les cartes, une analyse QTL par Interval Mapping a été réalisée. Cette analyse a été effectuée sur les groupes de liaison sur lesquels une EST avait été cartographiée, soit 3 groupes chez *E. urophylla* et 7 groupes chez *E. grandis*. Les résultats de cette analyse QTL sont présentés dans le tableau 8.

- Détection de QTL chez *E. grandis* :

Chez ce parent, 17 QTL ont été détectés par cette méthode. Pour le taux de lignine de Klason, 4 QTL ont été détectés sur les groupes de liaison 4, 10 et 11 (ce dernier en ayant 2). Les caractères chimiques permettent de mettre en évidence 10 QTL : 4 pour la lignine (GL4, 10 et 11), 3 pour S/G (GL6 et GL11), 1 pour S (GL2), 1 pour H+G+S (GL4) et 1 pour la teneur en flavanols (GL11). Le GL4 qui a un QTL de teneur en lignines a également un QTL pour la teneur en H+G+S. Sur les GL2, 3 et 6, des QTL pour le pilodyn ont été détectés. Pour la DRLM, 2 QTL ont été détectés sur les groupes de liaison 4 et 11. Pour la longueur moyenne des fibres et l'indice d'éclatement, 1 seul QTL a été mis en évidence, respectivement sur GL1 et GL4. Une seule colocalisation a été mise en évidence pour l'EST 6D07 avec un QTL de pilodyn à 51 mois. Pour les autres QTL, peu sont proches des EST cartographiées, les distances entre QTL et l'EST la plus proche variant de 3,3 à 61,3 cM.

- Détection de QTL chez *E. urophylla* :

Pour *E. urophylla*, sur les 3 GL qui ont une EST comme marqueur, 10 QTL ont été mis en évidence : 3 sur GL1 (2 sur GL1a et 1 sur GL1b), 5 sur GL2 et 2 sur GL9 (illustrés sur la figure 8). Les 2 QTL de GL1a concernent la teneur en G et le rapport S/G, mais à des positions différentes (respectivement 9 et 26 cM). Le QTL pour S/G colocalise avec l'EST 6A07. Le GL1b a un QTL pour la longueur

Tableau 9 : Résultats des tests de permutation (10000 permutations): A/ chez *E. grandis* et B/ chez *E. urophylla*. Le F max ainsi que le F à 1% sont fournis, de même que les F observés aux QTL et au marqueur le plus proche.

A/

Caractère	Groupe de liaison	Marqueur flanking	F obs. au QTL	F obs. au marqueur	F(1%)	Fmax
Klason	4	N19_1161/2+	12,67	8,38	6,86	19,62
Klason	11	I14_797/1+	9,43	10,28	6,55	17,67
Klason	10	A20_800/1-	11,26	9,37	6,91	15,14
Klason	11	X07_1583/1+	9,67	10,44	6,51	17,76
S(%boisext)	2	O06_1402/1-	10,30	7,21	6,76	17,11
H+G+S (µmole/glig)	4	E02_000/1-	9,48	6,01	6,73	13,88
S/G	6	X07_396/2+	15,21	12,95	6,73	16,43
S/G	6	K12_657/1-	11,26	11,10	6,87	16,14
S/G	11	R13_377/2+	14,77	13,17	6,85	16,35
estdma	11	A09_619/1+	18,46	17	6,64	15,91
DRLMmoy	4	N19_1161/2+	12,70	13,20	6,34	15,00
DRLMmoy	11	U20_1601/2-	11,83	9,97	7,01	15,41
Pil51sous	3	M05_1572/1+	13,97	14,41	7,13	17,03
Pil51sous	2	O06_1402/1-	10,19	13,63	6,79	14,18
Pil51sous	6	N11_419/1-	16,22	16,54	6,55	17,22
IEtot1sem	4	N19_1161/2+	9,72	6,94	7,01	14,99
longmoypond(mm)	1	E08_000/1+	16,62	15,57	6,86	16,71

B/

Caractère	Groupe de liaison	Marqueur flanking	F obs. au QTL	F obs. au marqueur	F(1%)	Fmax
G	1a	Y11_815/1+	17,84	15,31	6,87	19,55
S/G	1a	6A07_000/1+	11,02	4,91	6,85	21,73
	9	X12_383/1-	10,19	10,13	6,96	15,15
ESmoy30	2	A20_938/1+	19,49	17,29	6,69	14,96
	4b	B08_1945/2+	10,91	12,78	7,13	21,28
Pil59sous	1a	Z03_925/1+	10,38	10,69	6,58	16,93
	9	X09_1113/1-	9,52	9,88	6,93	16,24
longmoypond(mm)	1b	M06_1453	10,50	8,68	7,01	13,14
	2	6D01_000/1+	16,23	13,50	7,01	18,57
estdma	2	6D01_000/1+	12,58	5,80	6,84	19,50

moyenne pondérée des fibres de bois. Pour GL2, 1 QTL a été trouvé pour la teneur en phénols, 1 pour le pilodyn sous écorce à 51 mois et un pour la longueur moyenne pondérée des fibres de bois. Ces 3 QTL colocalisent avec l'EST 6D01. Un autre QTL pour le pilodyn a été détecté et il colocalise avec l'EST 6D12. Le 5ème QTL détecté sur ce groupe de liaison concerne le module d'élasticité spécifique à 30 mois; pour ce caractère, aucune colocalisation avec 1 EST n'a été mise en évidence. Pour GL9, 2 QTL pour des caractères impliqués dans la nervosité du bois (DRLM et IEtot1J) ont été détectés, ces 2 QTL étant distants de 4 cM.

3.2.4 Tests de permutations

Le test de permutation permet de valider l'existence de QTL détectés par IM en redistribuant au hasard les génotypes des individus pour un marqueur donné. Un nouveau F est calculé, au seuil de 5%, 1% et au maximum. Chez *E. urophylla*, 90% des QTL ont un F observé au marqueur le plus proche supérieur au F à 1% et parmi ces 90%, 22,2% ont un F supérieur au F max. Pour *E. grandis*, sur les 17 QTL détectés, 94,1% ont un F observé au marqueur supérieur au F à 1%, dont 6,25% un F observé supérieur au F max. Chez ces 2 parents, tous les QTL détectés par IM ont un F observé au QTL supérieur au F à 1% (tableau 9).

4 DISCUSSION

4.1 Cartographie génétique des EST

4.1.1 Détection du polymorphisme

La technique SSCP est une méthode très sensible car elle permet de détecter des mutations ponctuelles (Orita *et al.*, 1989 [41]). Cependant, dans notre étude, seuls 58,3% des fragments amplifiés se sont révélés polymorphes. La détection de mutations ponctuelles par SSCP est optimale pour des fragments d'une taille inférieure à 200 pb (Hayashi *et al.*, 1993 [28]). Or, sur les 10 EST monomorphes, 8 ont une taille supérieure. Ces fragments, s'ils présentent du polymorphisme dans la descendance étudiée, ce dernier ne semble pas détectable par SSCP. Proportionnellement, plus les fragments amplifiés sont grands, plus la variation de séquence doit être importante mais plus des mutations ponctuelles sont difficiles à détecter. Dans cette étude, on observe que sur les 6 fragments contenant un intron, 5 se sont révélés polymorphes. Plomion *et al.* (1999 [47]) ont montré que pour 16 séquences de gènes de pin maritime, la présence d'intron (pour 8 séquences) n'améliorait pas la détection de polymorphisme par SSCP. Ces résultats SSCP ne traduisent pas le niveau de polymorphisme plus important généralement rencontré dans les régions non codantes. En effet, Champurney (2003 [12]) a montré que les variations nucléotidiques au sein du gène CCR pour 40 génotypes *E. urophylla* étaient principalement situées dans les zones introniques, ce qui pourrait expliquer nos résultats. L'absence de détection de polymorphisme dans un fragment ne signifie pas pour autant une absence de variation dans le gène. En effet, les amorces utilisées peuvent ne pas cibler le ou les sites polymorphes. Pour la suite des travaux, il serait donc intéressant de choisir les amorces de telle manière qu'elles amplifient un intron. Une autre alternative serait de séquencer les EST chez les parents du croisement avant de faire les tests SSCP pour s'assurer qu'une variation nucléotidique existe. Enfin, un microsatellite a été mis en évidence pour l'EST 6A11 (données non présentées). Les microsatellites étant des locus hautement polymorphes, ce type de marqueur pourrait également faciliter la cartographie des EST.

4.1.2 Conservation des séquences entre espèces

Les amorces utilisées pour amplifier les EST chez *E. globulus*, *E. urophylla* et *E. grandis* ont été définies sur des séquences spécifiques à *E. gunnii*. Elles ont permis d'obtenir 82,4% de produits exploitables, pour la SSCP, chez *E. globulus*, 73,5%

chez *E. grandis* et 70.6% chez *E. urophylla*. *E. globulus* et *E. gunnii* appartiennent à la section *Maidenaria* alors qu'*E. urophylla* et *E. grandis* appartiennent à la section *Transversaria*, indiquant que les séquences semblent mieux conservées dans une même section. Gion (2001 [21]) a également utilisé des séquences d'*E. globulus* et d'*E. gunnii* pour amplifier des gènes chez *E. urophylla* et *E. grandis*. Après séquençage, une homologie de séquence supérieure à 80% a été trouvée entre les séquences amplifiées et la séquence des gènes utilisés pour la définition des amorces. Champurney (2003 [12]) a observé le même type de résultats pour le gène CCR d'*E. urophylla*, d'*E. gunnii*, d'*E. saligna* et d'*E. globulus*. De plus, Brondani *et al.* (2002 [7]) ont montré que 90% des microsatellites développés étaient transférables dans une même section et plus de 80% le sont au sein du sous-genre *Symphyomyrtus*. Ces différents résultats révèlent une bonne conservation des génomes entre les espèces d'*Eucalyptus* appartenant au sous-genre *Symphyomyrtus*.

4.1.3 Informations apportées par la cartographie d'EST

Les travaux de cartographie réalisés dans cette étude ont permis de confirmer que les cartes génétiques des deux parents du croisement étudié ne sont pas totalement saturées. Cette conclusion avait déjà été évoquée par Gion (2001 [21]) qui avait déterminé que les cartes génétiques générées couvrent 80% du génome d'*E. grandis* et 93% du génome d'*E. urophylla*. La cartographie d'EST surexprimées dans le xylème, sur différents groupes de liaison d'*E. urophylla* et d'*E. grandis*, n'a pas révélé de cluster de gènes. Cette indépendance entre locus est favorable à leur utilisation en SAM. En effet, la probabilité qu'un génotype combinant les allèles favorables aux meilleurs gènes existe, est plus importante.

Jusqu'à récemment, les marqueurs dominants, et principalement les RAPD, ont permis d'établir des cartes génétiques chez quelques espèces d'eucalyptus (Grattapaglia *et al.*, 1994 [25]; Verhaegen et Plomion 1996 [57]). Cependant, ces marqueurs ne permettent pas d'étudier la sythénie entre différentes espèces. De ce fait, les marqueurs codominants sont de plus en plus utilisés. La codominance des marqueurs SSCP nous a permis de confirmer des homologies entre groupes de liaison des deux espèces. Des régions homologues ont aussi été délimitées grâce aux EST polymorphes chez les deux parents. La conservation des génomes entre espèces permet d'envisager le positionnement des ces gènes de fonction connue sur les cartes génétiques d'autres espèces d'*Eucalyptus*. L'ajout de gènes candidats sur ces cartes devrait permettre à

long terme des études de cartographie comparée (synthénie) et devrait constituer de bon outils pour des analyses de diversité.

4.2 Détection de QTL

4.2.1 Détection de QTL marqueur par marqueur *vs* détection par IM

La comparaison des résultats de détection de QTL obtenus entre les méthodes marqueur par marqueur et IM montre des différences dans le nombre et les effets des QTL mis en évidence. En effet, on s'aperçoit qu'un QTL peut être détecté par une méthode mais pas par l'autre. Une première raison vient du type d'analyse qui est réalisé. En effet, la détection de QTL par IM est réalisé sur un intervalle (en cM) en prenant en compte le pourcentage de recombinaison entre les marqueurs bordants. De ce fait, des QTL détectables sur cet intervalle peuvent ne pas être détectés au niveau du marqueur. Une autre explication peut être l'effectif utilisé pour chaque analyse. En effet, par ANOVA les calculs sont réalisés sur l'effectif réel disponible pour les EST, soit 92 individus. Par IM, il sont réalisés sur l'ensemble de la descendance (201 individus) grâce à une fonction permettant de compléter les données moléculaires manquantes des marqueurs en se basant sur les information fournies par les marqueurs adjacents. La précision du test est ainsi augmentée grâce à un nombre plus important d'individus.

4.2.2 Absence de colocalisation

Pour 10 EST, aucune colocalisation n'a été mise en évidence. Parmi ces EST, certaines codent les enzymes intervenant dans les voies métaboliques de la formation des parois, comme l'EST 7A07 codant pour une xyloglucane endotransglycosylase, ou encore l'EST 11H04 codant pour une enzyme intervenant dans la voie de biosynthèse de la cellulose. La non-colocalisation de ces EST avec un QTL ne signifie pas qu'elles n'aient pas un effet important pour les caractères quantitatifs étudiés. Cependant, il semble que dans le croisement étudié, leur variabilité n'ait pas d'effet sur la variation observée. La différence d'effet entre les allèles mis en évidence n'est pas suffisante pour induire une variation phénotypique. Ces gènes pouvant être de bons candidats dans d'autres fonds génétiques, il pourrait être intéressant de réaliser le même type d'étude à partir d'autres croisements et/ou pour d'autres espèces d'eucalyptus. De plus, ces enzymes qui ont des rôles importants dans la xylogénèse, ne sont pas directement liées aux caractères mesurés. Elles jouent un rôle en amont

et les caractères chimiques mesurés sont des produits finaux des voies de biosynthèse (lignines de Klason, monomères de la lignine). Ceci peut aussi expliquer l'absence de colocalisation EST/QTL.

4.2.3 Colocalisation des EST avec des QTL de qualité du bois

La détection de QTL a permis de mettre en évidence des colocalisations entre les EST 6D01, 6D07, 6D12 et 7A09 (EST non positionnée sur les cartes génétiques) et des QTL de qualité du bois. Les EST 6D07 et 6D01 pour lesquelles la fonction n'est pas connue (aucune similarité avec des gènes de fonction connue dans les banques) semblent correspondre à des gènes importants pour la xylogénèse. En effet, leur niveau d'expression dans le xylème d'*E. gunnii* par rapport à la feuille est élevé (respectivement 2,9 et 4,2). De plus, leurs colocalisations avec des QTL de qualité du bois chez *E. urophylla* et *E. grandis*, semblent montrer que la variabilité moléculaire de ces gènes est responsable d'une partie de la variation de plusieurs propriétés mécaniques, anatomiques, chimiques du bois. Des études plus poussées pour connaître la fonction de ces 2 EST devront être menées. L'EST 7A09 quant à elle code pour une pectine acétyl estérase qui intervient dans le métabolisme des pectines. Sa colocalisation avec des teneurs en syringyl et en guaiacyl pourraient traduire un rôle de ce gène sur l'imprégnation des lignines dans les parois pectocellulosiques. Cette EST n'étant pas liée aux groupes de liaison d'*E. grandis* et d'*E. urophylla*, il serait nécessaire de saturer les cartes génétiques pour localiser cette EST sur le génome d'eucalyptus et préciser son effet sur la variation des teneurs en monomères syringyl et guaiacyl. Pour les EST qui colocalisent avec des QTL de qualité du bois, il sera malgré tout nécessaire de valider leurs effets en populations afin d'envisager leur utilisation en SAM. Ce type d'étude nécessite d'avoir accès aux différentes formes alléliques présentes dans la population étudiée grâce à la mise en évidence des variations nucléotidiques sur toute ou partie du gène (étude SNP : Single Nucleotide Polymorphism). De telles études ont été engagées au CIRAD pour le gène CCR sur *E. urophylla*. Les premiers résultats semblent encourageants pour élargir ce type d'étude à d'autres gènes (Champurney, 2003 [12]).

5 CONCLUSION et PERSPECTIVES

La cartographie génétique d'EST développées chez *E. gunnii* sur les cartes de deux autres espèces d'*Eucalyptus* a été efficace en utilisant la technique SSCP. Ces résultats montrent une bonne conservation des génomes entre espèces du sous genre *Symphyomyrtus* et semblent encourageants pour développer des approches de cartographie comparée entre génomes. Cependant, d'autres stratégies de mise en évidence de polymorphisme, sont à développer pour les EST supposées monomorphes pour lesquelles la sensibilité de la technique utilisée n'est peut être pas suffisante pour révéler l'ensemble des variations nucléotidiques.

Des colocalisations EST/QTL de qualité du bois ont été mises en évidence. Ces résultats sont prometteurs pour l'utilisation de ces gènes en SAM. Cependant, les cartes génétiques doivent être complétées et la détection de QTL doit se faire sur la totalité des individus disponibles pour préciser l'effet des QTL. Ces deux points sont des conditions indispensables pour que les résultats soient entièrement validés. De plus d'autres études sur les EST qui colocalisent semblent indispensables pour confirmer leur fonction, connue uniquement par homologie de séquences. L'accès à la séquence entière des gènes correspondant aux EST permettrait d'étudier par transformation génétique sur tabac ou arabidopsis (difficile chez l'eucalyptus), l'effet de sur ou sous expression de ces gènes sur différents caractères d'intérêts : par des approche anti-sens, ou d'inactivation de gènes par insertion de transposons (Pflieger *et al.*, 2001 [42]). Parallèlement à ces études fonctionnelles, des approches populationnelles (populations naturelles ou d'améliorations) pourraient être menées. En effet, la mise en évidence de polymorphisme SNP pour les EST d'intérêt permettrait de révéler les différentes formes alléliques présentes au sein des espèces utilisées et d'estimer leur effet respectif sur les caractères cibles de la sélection. Certains résultats récents ont montré que le parent *E. urophylla* du croisement étudié, présente une mutation non synonyme au sein du gène CCR (Champurney, 2003 [12]), gène qui colocalise avec un QTL de teneur en lignines chez ce parent.

Avant d'envisager une utilisation de gènes candidats en sélection assistée par marqueurs, de nombreux travaux restent à mener. En effet, notre étude n'est qu'une première étape dans la mise en évidence de variabilité fonctionnelle de gènes candidats.

Bibliographie

- [1] I. Allona, M. Quinn, E. Shoop, K. Swope, S. St. Cyr, J. Carlis, J. Riedl, E. Retzel, M.M. Campbell, R. Sederoff, and R.W. Whetten. Analysis of xylem formation in pine by cDNA sequencing. *Proc. Natl. Acad. Sci. Plant Biology*, 95:9693-9698, 1998.
- [2] H. Baillères. *Précontraintes de croissance et propriétés mécano-physiques de clones d'Eucalyptus (Pointe-Noire. Congo): hétérogénéités, corrélations et interprétations histologiques*. Thèse de doctorat, Université de Bordeaux I, 1994.
- [3] C. Bodénès, F. laigret, and A. Kremer. Inheritance and molecular variations of PCR-SSCP fragment in Pedunculate oak (*Quercus robur* L.). *Theor. Appl. Genet.*, 93:348-354, 1996.
- [4] J.M. Bouvet. *Evolution de la variabilité avec l'âge et corrélation juvenile-adulte dans des populations d'Eucalyptus*. PhD thesis, Institut National Agronomique Paris-Grignon, 1995.
- [5] H.D. Bradshaw and R.F. Stettler. Molecular genetics of growth and development in Populus IV. Mapping QTLs with large effects on growth, form and phenology traits in a forest tree. *Genetics*, 139:963-973, 1995.
- [6] H.D. Bradshaw, M. Villar, B.D. Watson, K.G. Otto, S. Stewart, and R.F. Stettler. Molecular genetics of growth and development in Populus III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. *Theor. Appl. Genet.*, 89:167-178, 1994.
- [7] R.P.V. Brondani, C. Brondani, and D. Grattapaglia. Towards a genus-wide reference linkage map for Eucalyptus based exclusively on highly informative microsatellite markers. *Mol. Genet. Gen.*, 267, 2002.
- [8] R.P.V Brondani, C. Brondani, R. Tarchini, and D. Grattapaglia. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theor. Appl. Genet.*, 97, 1998.
- [9] M. Byrne, J.C. Murrell, J.V. Owen, B. Allen, and G.F. Moran. An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. *Theor. Appl. Genet.*, 91, 1995.
- [10] M. Byrne, J.C. Murrell, J.V. Owen, P. Kriedemann, Williams E.R., and G.F. Moran. Identification and mode of action of quantitative trait loci affecting seedling height and leaf area in *Eucalyptus nitens*. *Theor. Appl. Genet.*, 94, 1997.

- [11] M.T. Cervera, V. Storme, Liu B., J. Gusmao, M. Steenackers, B. Ivens, B. Michiels, M. Van Montagu, and W. Boerjan. Inheritance of AFLP markers in *Populus*: genome mapping of *P. deltoides*, *P. nigra* and *P. trichocarpa*. In *Second International Wood Biotechnology Symposium: Proceedings A. Jamieson (Ed.)*, Ridgley, North Forest Products, 1999.
- [12] C. Champurney. *Mise en évidence de polymorphisme au sein du gène CCR chez Eucalyptus urophylla*. Rapport de DEA, Université de Montpellier II, 2003.
- [13] G.A. Churchill and R.W. Doerge. Empirical threshold values for quantitative trait mapping. *Genetics*, 138:963–971, 1994.
- [14] S. Creste, A. Tulmann Neto, and A. Figueira. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. *Plant Mol. Biol. Report.*, 19:299–306, 2001.
- [15] L. Diatchenko, Yun-Fai Chris Lau., A.P. Campbell, A. Chenchik, F. Moqadam, B. Huang, S. Lukianov, K. Lukianov, N. Gurskaya, E.D. Sverdlov, and P.D. Siebert. Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Biochemistry*, 93:6025–6030, 1996.
- [16] Food and Agricultural Organisation. Situation des forêts du monde 2001. In *Eds Organisation des Nations Unies pour l'alimentation et l'agriculture*, Rome (Italie), 2001.
- [17] B.E. Frewen, T.H.H. Chen, G.T. Howe, J. Davis, A. Rohde, W. Boerjan, and H.D. Bradshaw. Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. *Genetics*, 154:837–845, 2000.
- [18] H. Fukuda. Xylogenesis: initiation, progression, and cell death. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 47:299–325, 1996.
- [19] S. Fukuoka, T. Inoue, A. Miyo, Zhong H.S. Monna, L., T. Sasaki, and Y. Minobe. Mapping of sequence-tagged sites in rice by single-strand conformation polymorphism. *DNA Res.*, 1:271–277, 1994.
- [20] L.J. Gibson and M.F. Ashby. *Cellular solids: structure and properties*. Cambridge Solid State Science Series, 1999.
- [21] J.M. Gion. *Etude de l'architecture génétique des caractères complexes chez l'eucalyptus: des marqueurs anonymes aux gènes candidats*. Thèse de doctorat, Université de Rennes I, 2001.
- [22] J.M. Gion, P. Rech, J. Grima-Pettenati, D. Verhaegen, and C. Plomion. Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes. *Mo-*

lecular breeding, 6, 2000.

- [23] D. Grattapaglia, F.L.G. Bertolucci, R. Penchel, and Sederoff R. Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. *Genetics*, 144:1205–1214, 1996.
- [24] D. Grattapaglia, F.L.G. Bertolucci, and R. Sederoff. Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. *Theor. Appl. Genet.*, 90:933–947, 1995.
- [25] D. Grattapaglia and R. Sederoff. Genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. *Genetics*, 137, 1994.
- [26] J. Grima-Pettenati and D. Goffner. Lignin genetic engineering revisited. *Plant Science*, 145:51–65, 1999.
- [27] L. Harvengt, M. Pâques, and P. Philipon. La multiplication des pins. *Biofutur*, 199:22–25, 2000.
- [28] K. Hayashi and D.W. Yandell. How Sensitive is PCR-SSCP? *Human Mutation*, 2:338–346, 1993.
- [29] M. Hertzberg, H. Aspeborg, J. Schrader, A. Andersson, R. Erlandsson, K. Blomqvist, R. Bhalerao, M. Uhlén, T.T. Teeri, J. Lundeberg, B. Sundberg, and P. Nilsson. A transcriptional roadmap to wood formation. *PNAS*, 98:14732–14737, 2001.
- [30] A. Kremer. Diversité génétique et variabilité des caractères phénotypiques chez les arbres forestiers. *Genet. Sel. Evol.*, 26:105–123, 1994.
- [31] E.S. Lander, P. Green, J. Abrahamson, A. Barlow, M.J. Daly, and Newburg L. Lincoln, S.E. MARMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, 1:174–181, 1987.
- [32] C. Lapierre, C. Rolando, and B. Monties. Thioacidolyse of poplars lignins: identification of monomeric syringyl products and characterization of guaiacyl-syringyl lignins fractions. *Holzforschung*, 40:1215–1225, 1986.
- [33] G. Leprovost. *Effet de la saison et d'un stress mécanique sur la variation du transcriptome dans le xylème en formation chez le pin maritime (Pinus pinaster Ait.)*. Thèse de doctorat, Université de Bordeaux I, 2003.

- [34] D. Maillard. La pollinisation contrôlée (eucalyptus, pins). *Document interne du Centre Technique Forestier Tropical du Congo*, page 54 p., 1978.
- [35] H. Mariaux. Structure fine du bois de wapa en relation avec les contraintes de croissance. In *Sciences et industries du bois groupe 1 (colloque)*, 1982.
- [36] B. Martin and G. Quillet. Bouturage des arbres forestiers au Congo : résultats des essais effectués à Pointe-Noire de 1969 à 1973. *Bois et Forêts des Tropiques*, 154:41–57, 1974.
- [37] G.F. Moran and J.C. Bell. *Isozymes in Plant Genetics and Breeding, Part B*. Tanksley S.D. and Orton T.J. (Eds), 1983.
- [38] R.M. Myers, N. Lumelsky, L.S. Lerman, and T. Maniatis. Detection of single base substitutions in total genomic DNA. *Nature*, 313:495–497, 1985.
- [39] D. Neale, K. Jermstad, M. Sewell, and N. Wheeler. Progress towards detecting and verifying QTLs for wood property traits in loblolly pine and adaptive traits in Douglas-fir. In *Proceedings IUFRO Conference*, 1997.
- [40] J.C. Nelson. QGENE: software for marker-based genomic analysis and breeding. *Mol. Breed.*, 3:239–245, 1997.
- [41] M. Orita, H. Iwahana, H. Kanazawa, K. Hayashi, and T. Sekiya. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Genetics*, 86:2766–2770, 1989.
- [42] S. Pflieger, V. Lefebvre, and M. Causse. The candidate gene approach in plant genetics : a review. *Molecular Breeding*, 7:275–291, 2001.
- [43] C. Plomion. *Développement de la génomique pour l'étude de la variation des caractères d'intérêt chez le pin maritime*. Mémoire d'habilitation à diriger des recherches, Bordeaux, 2000.
- [44] C. Plomion, N. Bahrman, and O'Malley D. Durel, C.E. Genomic analysis in *Pinus pinaster* (maritime pine) using RAPD and protein markers. *Heredity*, 74:661–668, 1995.
- [45] C. Plomion and C.E. Durel. Estimation of the average effects of specific alleles detected by the pseudo-testcross QTL mapping strategy. *Gen. Sel. Evol.*, 28, 1996.
- [46] C. Plomion, C.E. Durel, and D. O'Malley. Genetic dissection of height in maritime pine seedlings raised under accelerated growth condition. *Theor. Appl. Genet.*, 93, 1996.
- [47] C. Plomion, P. Hurme, J.M. Frigerio, M. Ridolfi, D. Pot, C. Pionneau, C. Avila, F. Gallardo, H. David, G. Neutelings, M. Campbell, F.M. Canovas, O. Savo-

- lainen, C. Bodenes, and A. Kremer. Developping SSCP markers in two Pinus species. *Mol. Breeding*, 5:21–31, 1999.
- [48] C. Plomion, G. Leprovost, and A. Stokes. Wood Formation in Trees. *Plant Physiol.*, 127:1513–1523, 2001.
- [49] M. Quint, R. Mihaljevic, C.M. Dussle, M.L. XU, A.E. Melchinger, and T. Lübberstedt. Development of RGA-CAPS markers and genetic mapping of candidate genes for sugarcane mosaic resistance in maize. *Theor. Appl. Genet.*, 105:355–363, 2002.
- [50] D.G. Riesner, U. Steger, M. Wiese, M. Wulfert, M. Heibey, and K. Henco. Temperature-gradient electrophoresis for the detection of polymorphic DNA and for quantitative polymerase chain reaction. *Electrophoresis*, 13, 1992.
- [51] W. Rychlik, W.J. Spencer, and R.E Rhoads. Optimization of the annealing temperature for DNA amplification in vitro. *Nucl. Acids Res.*, 18:6409–6412, 1990.
- [52] M.A. Saghai-Marooof, K.M. Soliman, R.A. Jorgensen, and R.W. Allard. Ribosomal DNA spacer-length polymorphisms in barley : Mendelian inheritance, chromosomal location, and population dynamics. *proc. Natl. Acad. Sci.*, 81:8014–8018, 1984.
- [53] F. Sterky, S. Regan, M. Rohde A. Karlsson, J. Hertzberg, A. Holmberg, B. Amini, R. Bhalerao, M. Larsson, R. Villarroel, M. Van Montagu, G. Sandberg, O. Olsson, T.T. Teeri, W. Boerjan, P. Gustafsson, M. Uhlén, B. Sundberg, and J. Lundeberg. Gene discovery in the wood-forming tissues of poplar : Analysis of 5692 expressed sequence tags. *Proc. Natl. Acad. Sci. Plant Biology*, 95:13330–13335, 1998.
- [54] H. Sävenstrand, M. Brosché, M. Ängelsten, and Å. Strid. Molecular markers for ozone stress isolated by suppression subtractive hybridization: specificity of gene expression and identification of a novel stress-regulated gene. *Plant, Cell and Environment*, 23:689–700, 2000.
- [55] K.A. Thamarus, K. Groom, J. Murrell, M. Byrne, and G.F. Morane. A genetic linkage map for Eucalyptus globulus with candidate loci for wood fibre, and floral traits. *Theor. Appl. Genet.*, 104:379–387, 2002.
- [56] S. Tragoonrun, V. Kanizhin, P.M. Hayes, and T.K. Blake. Sequence tagged-site-facilitated PCR for barley genome mapping. *Theor. Appl. Genet.*, 84, 1992.
- [57] D. Verhaegen and C. Plomion. Genetic mapping in Eucalyptus urophylla and E grandis. RAPD markers. *Genome*, 39:1051–1061, 1996.

- [58] D. Verhaegen, C. Plomion, J.M. Gion, M. Poitel, P. Costa, and A. Kremer. Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *Theor. Appl. Genet.*, 95:597–608, 1997.
- [59] P. Vigneron. Création et amélioration de variétés d'hybrides d'*Eucalyptus* au Congo. In *IUFRO Symposium Intensive Forestry: the role of Eucalyptus*, Durban, South Africa, 1991.
- [60] P. Vigneron and J.M. Bouvet. *Amélioration des plantes tropicales; les eucalyptus*. Editions CIRAD, 1997.

RÉSUMÉ

Dans le cadre du programme d'amélioration génétique des eucalyptus au Congo, des travaux de cartographie génétique et de détection de QTL (Quantitative Trait Loci) pour la qualité du bois ont été entrepris depuis plusieurs années. Une approche gène candidat a récemment été mise en place avec des résultats encourageants pour les gènes de la voie de biosynthèse des lignines. L'objectif de cette étude est de positionner des étiquettes de gènes exprimés (EST) sur les cartes génétiques d'*Eucalyptus urophylla* et *E. grandis* déjà existantes et de tester l'association entre leur variabilité moléculaire et la variation de caractères de qualité du bois. Pour 34 EST d'*E. gunnii*, des amorces spécifiques ont été définies, et après optimisation des conditions d'amplification, la détection du polymorphisme au sein des fragments exploitables a été réalisée par SSCP (Single Strand Conformation Polymorphism). Au total, 14 EST polymorphes ont été cartographiées et une recherche de QTL a été effectuée. Parmi ces EST, 8 colocalisations avec des QTL de qualité du bois ont été mises en évidence. Même si ces résultats demandent à être précisés (effectif plus important pour les analyses) et validés au niveau des populations d'amélioration, ces résultats sont prometteurs pour envisager une sélection assistée par marqueurs pour la qualité du bois.

Mots clefs: *Eucalyptus*, gènes candidats, QTL, SSCP, qualité du bois

ABSTRACT

Genetic mapping and quantitative trait loci (QTL) analyses for wood quality have been conducted within the framework of the eucalyptus breeding programme in Congo. Recently, a candidate gene approach has been developed and good results were obtained for genes involved in lignin biosynthesis. The aim of the current study is first to map expressed sequence tags (EST) on the genetic maps of *Eucalyptus urophylla* and *E. grandis*, then to test effects of the genetic variability of candidate genes on the wood quality trait variation. Specific primers were designed for 34 EST of *E. gunnii* and after optimizing amplification conditions, the detection of polymorphism was realized using the Single Strand Conformation Polymorphism (SSCP) technic. A number of 14 EST were mapped and used for QTL analyses. Finally, 8 colocalisations between wood quality traits and EST were found. Even if these results need to be precised (analysis of larger sample) and validated by the study of gene effects within breeding populations, they open up new perspectives for markers assisted selection for wood quality.

Key words: *Eucalyptus*, candidate genes, QTL, SSCP, wood quality